PRIMARY PREVENTION AND EVALUATION RESOURCE KIT

Volume 3:

Analyzing Evaluation Data

PENNSYLVANIA COALITION AGAINST RAPE

This resource was supported with funding from the Pennsylvania Department of Health. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Pennsylvania Department of Public Health.

Recommended citation:

Townsend, S. (2014). *Primary prevention and evaluation resource kit* (Vol. 3). Enola, PA: Pennsylvania Coalition Against Rape

TABLE OF CONTENTS

OVERVIEW

Spirit of Evaluation

In the movement to end sexual violence, questions of effectiveness are especially important because the stakes are high.

In almost all areas of social services, community development and education the question is being asked: What impact does your program have? This question is often asked by funders who want to ensure they are being good stewards of the resources they manage. However, it is also a question that many program staff ask. Survivors depend on us to provide services that will help them heal. Parents depend on us to help them keep their children safe. Communities depend on our leadership to prevent sexual violence. **So we need to know with greater certainty what we are achieving**. With that knowledge, we can make better informed decisions about our programs.

Resource Kit

This manual addresses the area of program evaluation that many agencies struggle with the most: quantitative data analysis. It is the third in the four-volume *Primary Prevention and Evaluation Resource Kit* and is a project of the Pennsylvania Coalition Against Rape. The four volumes are most effective when used together. However, they can be used independent of one another. Additionally, while the examples in this volume focus on primary prevention programs, the techniques of data management and analysis can be used for analyzing outcomes of all sorts. For example, the same skills taught here can be used to evaluate quantitative measures of counseling and advocacy services, public awareness campaigns, community readiness, etc.

This volume has four main sections:

SECTION 1:
Basics of Program Evaluation

SECTION 2:
Managing Evaluation Data

SECTION 3:
Analyzing Quantitative Evaluation Data

SECTION 4:
Interpreting Quantitative Results

The first section provides a brief review of program evaluation. Drawing from concepts that were described in depth in Volume 2, this review will address the main ideas of program evaluation. It will set the stage for the main focus of this manual: quantitative data analysis. The second section teaches basic skills for turning survey or other quantitative data into a database. The third section introduces the reader to techniques for analyzing quantitative data. The fourth section explores how to interpret quantitative results by translating the numbers that are calculated into meaningful conclusions about the program's effects.

Qualities of a Data Analyst

It is likely that every agency has a number of staff and/or volunteers who have the qualities needed to analyze quantitative data effectively. This does not require any specific academic background. This volume is written with the novice in mind. **It does not presume any background in statistics or mathematics beyond simple arithmetic.** Special care has been taken to explain concepts and procedures in plain language.

The procedures explained in this manual also do not require any special statistical software. Given the popularity of Microsoft Office, all instructions will be on how to use Microsoft Excel as a database. However, the same principles can be applied to any other spreadsheet software.

Who is likely to work well with quantitative data? Someone who:

- Likes solving puzzles
- Is detail oriented
- Can step back from the details to see the big picture
- Is good at prioritizing tasks and goals
- Wants to develop new skills
- Is comfortable with basic math such as counting, percentages and averages

Whether your agency evaluates its work on its own or contracts with an independent evaluator, the more staff understand the logic and mechanics of program evaluation, the better equipped they will be to make evidence-informed decisions. Therefore, this volume is designed to provide a basic understanding and fundamental skills that staff can use to interpret and use quantitative evaluation findings, work effectively with outside consultants and, if desired, carry out their own analysis of quantitative evaluation data.

BASICS OF PROGRAM EVALUATION

- Why Evaluate Program Outcomes?
- Types of Evidence
- Basic Steps of Program Evaluation
- Case Example

WHY EVALUATE PROGRAM OUTCOMES

As explained in detail in Volume 2, there are four common reasons for evaluating any type of program:

- **Evaluation can help program staff make informed decisions about continuing or modifying a program.** Evaluations can be used to identify programs that show promise or that demonstrate clear success, to identify programs that are not showing sufficient impact and to reveal whether the program is having undesirable effects. By providing a basis for informed decisions, evaluation protects programs from making capricious decisions.
- **Evaluation can help defend a program against outside criticism.** It provides checks and balances so that other factors (political climate, personal preferences, etc.) do not lead to arbitrary decisions about the continuation or cancellation of a program.
- **Evaluation can provide insight into how or why a program is or is not working.** Understanding the mechanisms by which programs work provides a foundation that staff can use when developing new programs.
- **Evaluation is a mechanism for accountability.** Public funds are limited. Citizens are entitled to know that their tax dollars are being used wisely. Private funders similarly want to know that their dollars are having a positive impact.

Often when people think about evaluation they have in mind something like the Consumer Reports. They want to rate programs so that they know which ones work or which ones are best. They may think of evaluation like a report card: How well did this program do as measured on a scale that is believed to be objective? However, evaluation is actually a much broader concept. Different types of evaluation answer different questions.

"Program evaluation" can refer to needs assessments, spelling out program theory, process evaluations, assessment of program outcomes and analyses of efficiency or cost-benefit analyses. This manual will focus on the assessment of **program outcomes**.

Evaluation is about more than merely rating programs or determining if a program worked.

TYPES OF EVIDENCE

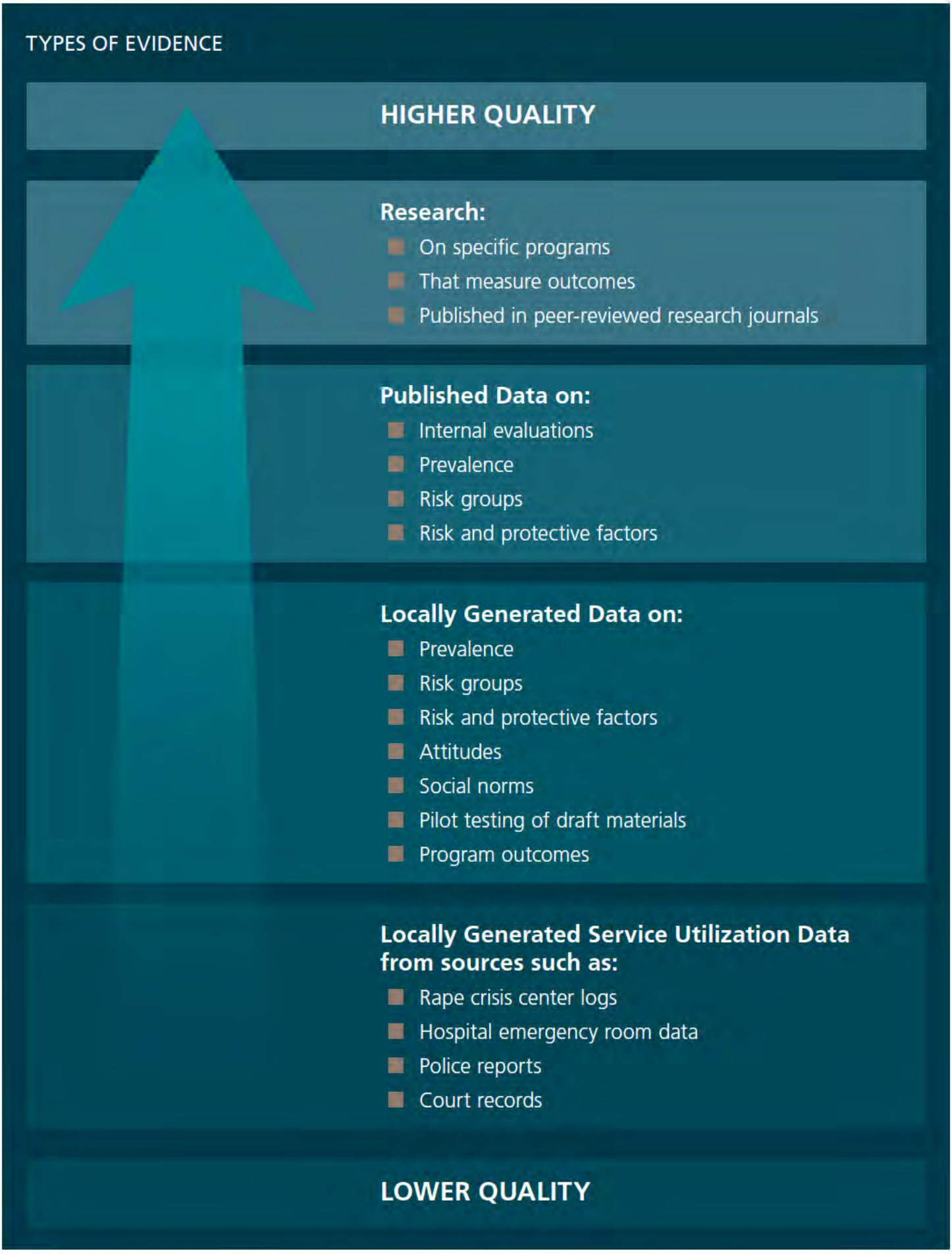
Evidence-Based Practice

One phrase that is often used is **evidence-based practice**. This refers to using strategies, curricula and campaigns that are based on sound evaluation or research evidence. Known in public health as efficacy trials (Glasgow, Lichtenstein, & Marcus, 2003), evidence-based practices are those that are based on trials where there was a high degree of control in the assessment, including delivering the program in a standardized and uniform way to a carefully selected audience. Due to the very strict conditions under which the program is delivered and its outcomes measured, a change is attributed to the preventive intervention.

This approach to evidence pictures different types of evidence as occurring on a continuum from *higher quality evidence* to *lower quality evidence* (Greco & Dawgert, 2007). (See figure on following page) Some funders are requiring the use of evidence-based practices that are supported by the highest quality research.

Requiring that grantees use strategies and programs that have been shown to have measurable and positive outcomes is an act of responsible public policy. Unlike the prevention of substance abuse or HIV/AIDS where "best practices" have begun to be identified, we do not yet know what the best practices are for preventing sexual violence. However, it is still important that we consider the evidence that is available.

Evidence-based practice presents particular challenges for sexual violence prevention because so little research on it has been done.


TYPES OF EVIDENCE
HIGHER QUALITY
Research:
On specific programs
That measure outcomes
Published in peer-reviewed research journals
Published Data on:
Internal evaluations
Prevalence
Risk groups
Risk and protective factors
Locally Generated Data on:
Prevalence
Risk groups
Risk and protective factors
Attitudes
Social norms
Pilot testing of draft materials
Program outcomes
Locally Generated Service Utilization Data from sources such as:
Rape crisis center logs
Hospital emergency room data
Police reports
Court records
LOWER QUALITY

Practice-Based Evidence

Practice-based evidence comes from real-life, community settings and reflects the values and practices in the movement.

In contrast, the kind of evaluation supported by this four-volume Resource Kit supports **practice-based evidence.** This type of evidence comes from real-life, community settings and reflects the values and practices in the movement.

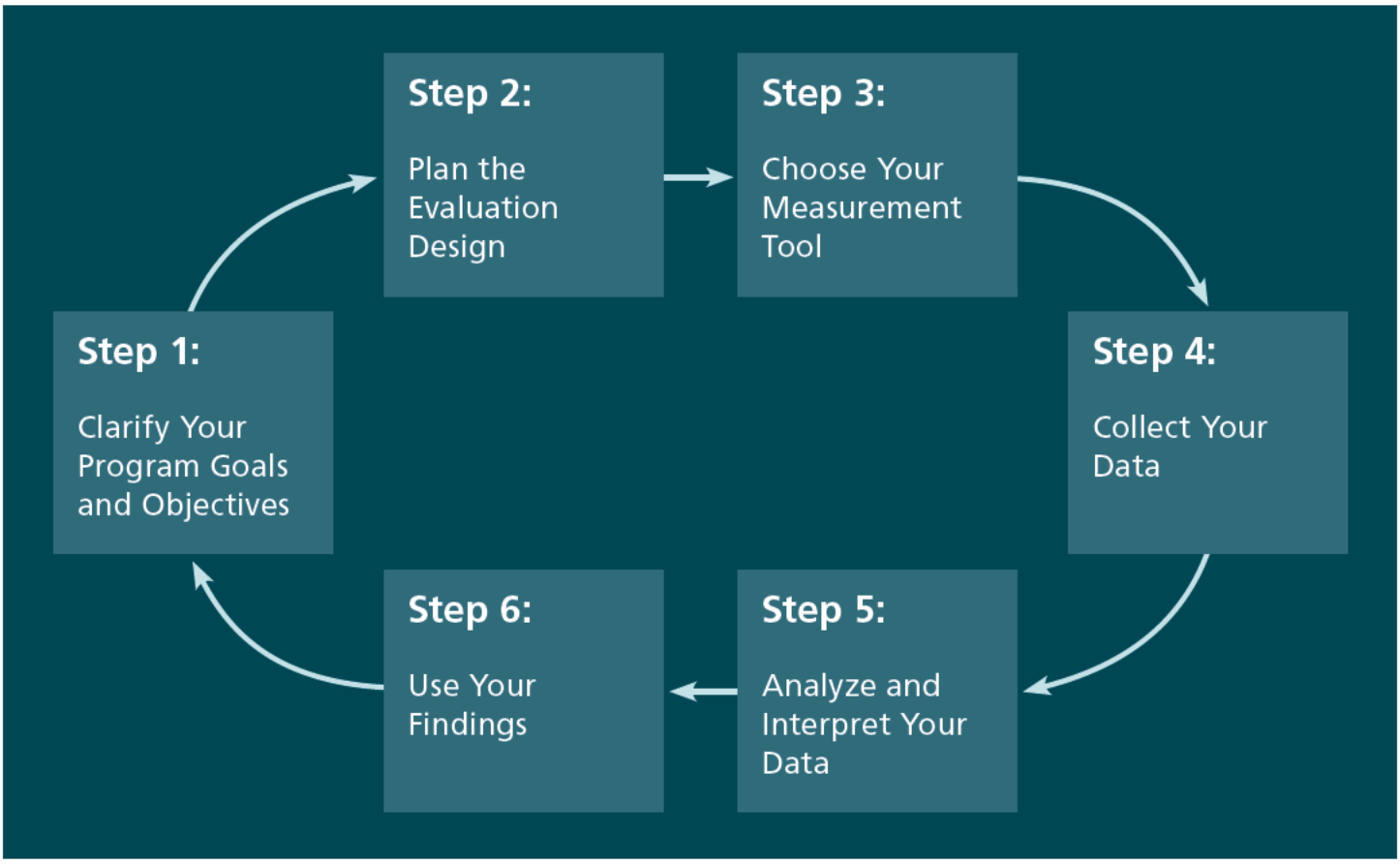
Practice-based evidence helps to identify what strategies can (Glasgow, Lichtenstein, & Marcus, 2003):

- Reach large audiences.
- Be used in different community settings.
- Be consistently implemented by facilitators.
- Produce sufficient effects using available resources.

Practice-based evidence often reflects a more relevant, useful and authentic expression of the practices used in the field. Practice-based evidence can only be generated through the leadership and involvement of the field. It may also be necessary to educate potential funders about the importance and validity of this type of evidence.

BASIC STEPS OF PROGRAM EVALUATION

Evaluating program impact has six basic steps:



Steps 1–4 are detailed in Volume 2. In summary:

Step 1: Clarify Your Program Goals and Objectives

The first step in any evaluation of program impact is to clarify the program's goals and objectives. This involves two tasks:

- Task 1: Describe the changes that should occur as a result of your program.
- Task 2: Based on that theory, define the program's goals and objectives.

Step 2: Plan the Evaluation Design

Planning your evaluation design involves three tasks:

- Task 1: Understand the design.
- Task 2: Determine the timing of your evaluation.
- Task 3: Determine who will participate in the evaluation.

The evaluation design that will be used in this manual is a Pre-Post Design with Follow up. The design looks like this:

Step 3: Select Your Measurement Tools

Once you have figured out the design for your evaluation you need to select the actual tools you will use to measure effectiveness. There are two main tasks in this step:

- Task 1: Select the type of measure you want to use.
- Task 2: Select the specific measure you will use and modify it as needed OR create your own measure.
- Task 3: Decide if answers will be anonymous or confidential.

The type of measure that will be used as an example of quantitative data in this manual is a written survey. However, the techniques described apply to any type of measure where the answers either are numbers or can be converted to numbers.

Step 4: Collect Your Data

Before you collect your data you should think carefully about what you are going to do. The goal is to ensure that all of the information is collected in a similar manner. This is true whether you are using surveys, focus groups, interviews or observations. You want to eliminate any variations that could influence your findings. The specific steps used to collect your data will depend on the type of measures you are using.

Step 5: Analyze and Interpret Your Data & Step 6: Using Your Findings

These are the steps detailed in this volume. The remainder of the volume will walk you through how to do this.

CASE EXAMPLE

Throughout this volume we will use a case example to illustrate each step of the data analysis process. The example is a simple one. A thorough evaluation of the program would include more than the five survey items described here.

CASE EXAMPLE

Changing Attitudes About Gender*

Community Rape Crisis Center (CRCC) is running a teen leadership program. It is based on Paul Kivel's *Making the Peace,* a violence prevention program for helping high school students build safer schools, relationships and communities. The 15-session curriculum addresses the social roots of violence and injustice and has exercises on building safe and inclusive communities. It looks at multiple forms of oppression including sexism, racism, classism and heterosexism. The program aims to turn teens into empowered bystanders who take action when they witness oppressive or violent acts.

CRCC is using the curriculum during a one-week summer camp. They then follow up with monthly meetings during the school year. In addition to the awareness and skill-building activities, the students spend time talking about their own experiences with oppression and get feedback on their attempts to act as empowered bystanders. They also decided to develop a sociodrama program that they are doing with younger children at local community centers, churches and synagogues.

The leadership program has many goals. One goal is to change teens' attitudes about gender. CRCC is particularly interested in promoting equity in dating and heterosexual relationships. To measure how well they achieve this outcome, they have chosen to have students complete the "Dating and Heterosexual Relationships Subscale" of the Pacific Attitudes Toward Gender Scale developed by Tracy Vaillancourt and Campbell Leaper. Students will complete the scale before the camp, at the end of the camp and at the end of the school year. The subscale has five items, shown on the next page.

* This example is inspired by CAMPpeaceworks, a project of Berks County Women in Crisis.

DATING AND HETEROSEXUAL RELATIONSHIPS SUBSCALE

1. I believe a man should be expected to pay the expenses on a date with a woman.	Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Mostly Agree	Strongly Agree
2. I believe it is all right for a woman to take the first steps to start a relationship with a man.	Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Mostly Agree	Strongly Agree
3. I believe it should be equally unacceptable for women and men to have sex with casual acquaintances.[1]	Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Mostly Agree	Strongly Agree
4. I believe a woman should be careful not to appear smarter than the man she is dating.	Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Mostly Agree	Strongly Agree
5. I believe when men show special courtesies only to women (like holding open the door), it reinforces the stereotype that women are helpless.	Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Mostly Agree	Strongly Agree

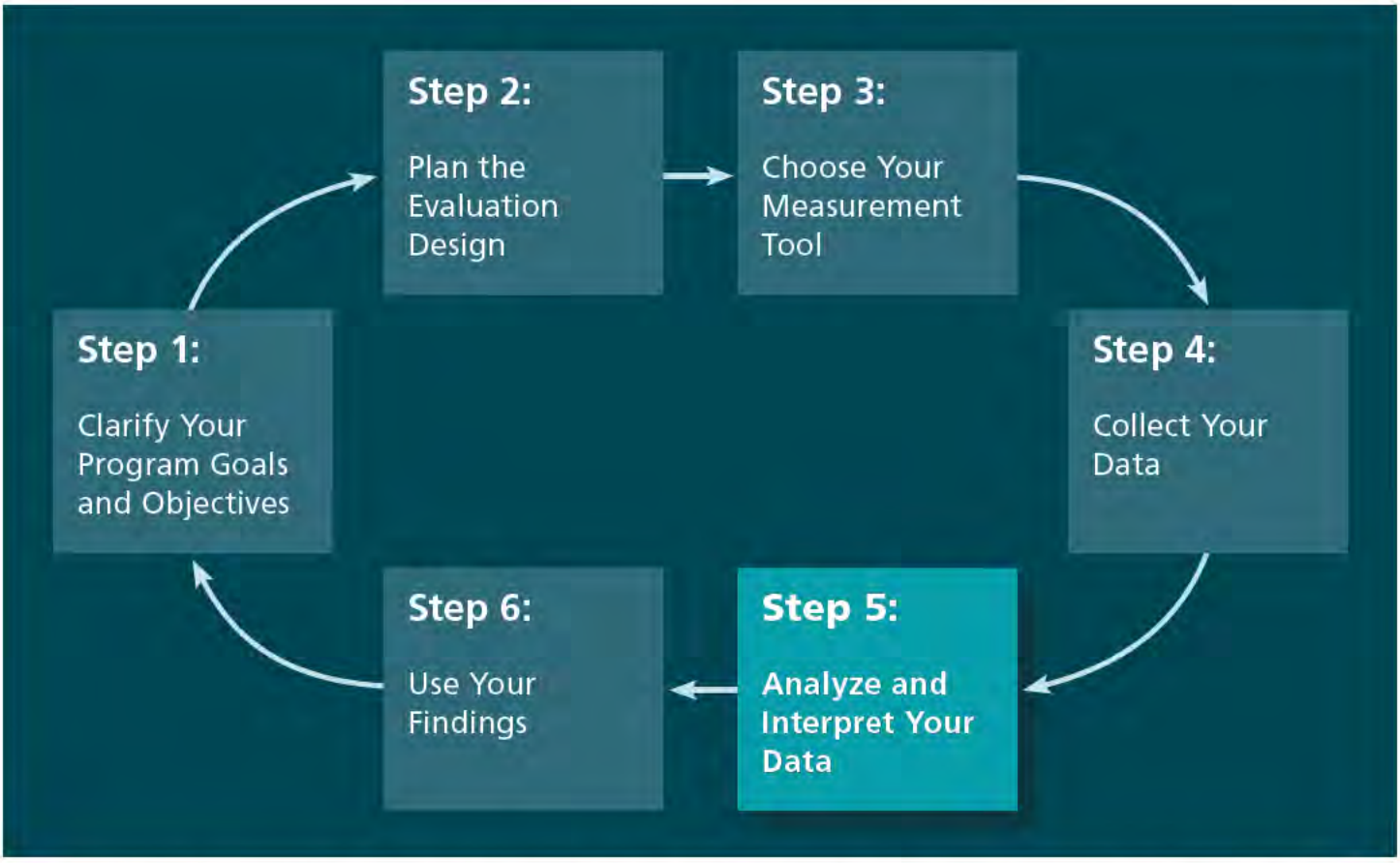
1 This item was modified from the original. The original measure states that it should be "equally acceptable" to have sex with casual acquaintances. However, because CRCC does not want to promote casual sex among teenagers they changed the item to say "equally unacceptable." This modification maintains the intent of the item by measuring whether women and men are being held to the same standard.

Throughout the remainder of this manual, these five questions will be used to illustrate how to analyze quantitative data.

MANAGING EVALUATION DATA

- Task 1: Create a data dictionary
- Task 2: Prepare survey/data forms
- Task 3: Set up your spreadsheet
- Task 4: Enter your data
- Task 5: Clean your data
- Task 6: Reverse code your data

MANAGING EVALUATION DATA



Once you have completed the first four steps of program evaluation, the next question is: *What do we do with all these surveys?* Preventionists may glance over the surveys as they come in and then put them in file cabinets until someone decides to clean house and recycle them all. Consequently, **many programs are sitting on a gold mine of data** that could be used to improve their programs, inform strategic planning, promote their programs in the community and support applications for funding.

To make use of the treasure trove in your file cabinets, you first need a system to help you organize and store your data. If designed appropriately, that system will let you move to the next stage of analyzing the data in meaningful ways.

The most common way to organize and store data (as well as to analyze it) is to put the survey responses into a computer spreadsheet. There are many options for spreadsheet software, but the one that your agency is likely to have already is Microsoft Excel.[2] The how-to instructions and pictures in this manual are for Microsoft Excel. If you are using a different program, the same basic principles should apply. Different versions of Excel may have slightly different features or appearances, but the illustrations in this manual should be very similar to what you see on your computer screen.

Note: The following instructions apply only to data that are based on closed-ended survey questions or archival data. They do not apply to open-ended survey questions, interviews or observations that rely on qualitative data.

To manage your data you need to follow six basic tasks:

Task 1: Create a data dictionary

Task 2: Prepare survey/data forms

Task 3: Set up your spreadsheet

Task 4: Enter your data

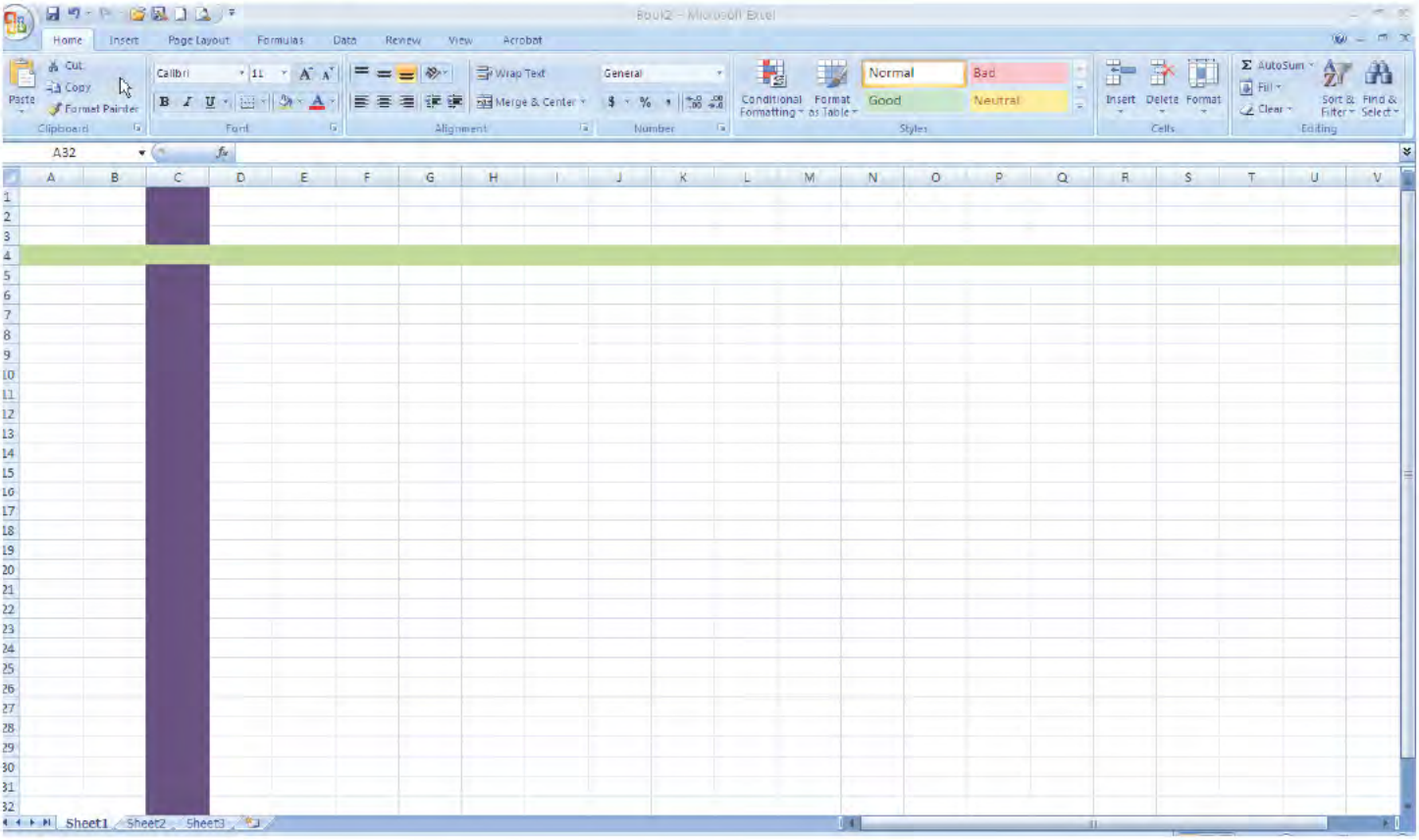
Task 5: Clean your data

Task 6: Reverse code your data (if necessary)

2 If your agency does not have the Microsoft Office suite and is not interested in purchasing it, you might want to use the free programs available through www.openoffice.org. The Calc program is the free equivalent of Microsoft Excel. Its appearance is very similar to Excel and it has the same basic functions.

TASK 1: CREATE A DATA DICTIONARY

The spreadsheet you will eventually create will be made of columns and rows. The picture below shows a column (Column C) highlighted in purple and a row (Row 4) highlighted in green. The boxes in the spreadsheet are called cells.



When you enter the data:

- You will have one **column** for each **question** on the survey.
- Each **person** who takes the survey will be a **row** in the spreadsheet.
- Their **answers** will be entered into the corresponding **cells**.
- All of the answers you put into the spreadsheet will need to be **numbers**.

A **data dictionary** will let you keep track of which column goes with which question and what the numbers mean.

A. Label Your Questions

The best way to create your dictionary is to take a **blank survey** and come up with a short label for each question. There are a few things to keep in mind when creating your labels:

Keep the labels **short**: 8–10 characters or less is best.

Each label must be **unique**, but because you will be asking the same questions on the Pre-test and Post-test you will want some **consistency** so you can match up the questions over time. You can either number the questions or use short words to refer to them; do whichever one will be easiest for you to keep track of.

Write the labels on the survey itself in a color of ink that will stand out.

At the top of that survey write the words **Data Dictionary** in big letters and save it in a folder where you or anyone else who enters data can easily find it.

Labels for the Dating and Heterosexual Relationships Subscale

The scale CRCC uses has five questions. The staff thought about two options for labeling the questions. The first option used the question numbers. It distinguished between pre-, post- and follow-up data by using the abbreviations "pre" "post" and "fu." Using this approach they came up with the following labels:

pre1	post1	fu1
pre2	post2	fu2
pre3	post3	fu3
pre4	post4	fu4
pre5	post5	fu5

They also thought about using the key word from each question, instead of numbers. Using this approach they came up with the following labels:

prepay	postpay	fupay
prefirst	postfirst	fufirst
presex	postsex	fusex
presmart	postsmart	fusmart
predoor	postdoor	fudoor

They decided to use key word labels. When the spreadsheet is looked at, they will automatically know to which question each label refers. If they used the number format they would either have to remember what Question 1, Question 2, etc. was about or they would have to look it up on the survey.

When CRCC wrote the labels on the copy of their survey it looked like this, where you see their handwritten labels in the margin on the right-hand side:

DATING AND HETEROSEXUAL RELATIONSHIPS SUBSCALE

prepay	1. I believe a man should be expected to pay the expenses on a date with a woman.	Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Mostly Agree	Strongly Agree
prefirst	2. I believe it is all right for a woman to take the first steps to start a relationship with a man.	Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Mostly Agree	Strongly Agree
presex	3. I believe it should be equally unacceptable for women and men to have sex with casual acquaintances.	Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Mostly Agree	Strongly Agree
presmart	4. I believe a woman should be careful not to appear smarter than the man she is dating.	Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Mostly Agree	Strongly Agree
predoor	5. I believe when men show special courtesies only to women (like holding open the door), it reinforces the stereotype that women are helpless.	Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Mostly Agree	Strongly Agree

They did the same thing on blank copies of their post- and follow-up surveys.

B. Code the Possible Answers

Notice that the answers students circle on the survey are words. **To do quantitative analysis we need to convert the words to numbers.** This is what we call **coding the data**. There are a few things to keep in mind when you are coding your data:

Make your codes logically increase in value from low to high where higher numbers indicate "more" of something. For example, in the scales below higher numbers mean "more agreement" or "more frequent" than lower numbers.

Be consistent. If you have two different scales with five points each, code both of them from 1 to 5. For example, you might have one scale that is about agreement and another that is about frequency. They both have five answers available.
Code them both from 1 to 5:

Agreement Scale	*Frequency Scale*
Strongly Disagree = 1	Never = 1
Disagree = 2	Rarely = 2
Neutral = 3	Sometimes = 3
Agree = 4	Often = 4
Strongly Agree = 5	Always = 5

Choose labels for categorical data. Some data do not increase logically. These are called categorical data and are not referred to as scales. For these choose an easy way of labeling the data. For example:

Male = 1
Female = 2

In this case a "2" is not "more" than a "1," but numbers are still used as codes.

Write the codes on your Data Dictionary for each survey.

Codes for the Dating and Heterosexual Relationships Subscale

The scale CRCC uses has six answers for students to show how much they agree with each statement. The staff decided to make **higher numbers** mean **more agreement** and **lower numbers** mean **less agreement**. They also asked students to identify their gender. Therefore, the codes they used were as follows:

Strongly Disagree = 1
Mostly Disagree = 2
Slightly Disagree = 3
Slightly Agree = 4
Mostly Agree = 5
Strongly Agree = 6

Male = 1
Female = 2

When CRCC wrote the codes on their data dictionary they looked like this, where you see their codes at the top of the columns:

		1	2	3	4	5	6
prepay	1. I believe a man should be expected to pay the expenses on a date with a woman.	Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Mostly Agree	Strongly Agree
prefirst	2. I believe it is all right for a woman to take the first steps to start a relationship with a man.	Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Mostly Agree	Strongly Agree
presex	3. I believe it should be equally unacceptable for women and men to have sex with casual acquaintances.	Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Mostly Agree	Strongly Agree
presmart	4. I believe a woman should be careful not to appear smarter than the man she is dating.	Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Mostly Agree	Strongly Agree
predoor	5. I believe when men show special courtesies only to women (like holding open the door), it reinforces the stereotype that women are helpless.	Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Mostly Agree	Strongly Agree

C. Decide on Rules for Coding Missing and Unclear Data

In a perfect world, your codes would be sufficient. However, **survey data are never perfect**. People will:

- Skip questions (intentionally or unintentionally).
- Give unclear answers.
- Make up their own answers.

You need to come up with **rules** for coding missing and ambiguous data that are logical so you can apply them **consistently** when it comes to entering the data. Examples of some common rules are:

- Any **missing** answers (i.e., questions a person skipped) will be labeled with a symbol (e.g., ~).
- Any answers that indicate the person's "real" answer is **in-between** two numbers will be coded with the lower value **OR** they will be entered as missing. Note: **Chose one of these rules and use the same rule every time.**

Examples of In-Between Answers for the Dating and Heterosexual Relationships Subscale

1. I believe a man should be expected to pay the expenses on a date with a woman.	Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Mostly Agree	Strongly Agree
2. I believe it is all right for a woman to take the first steps to start a relationship with a man.	Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Mostly Agree	Strongly Agree

- Any **invalid** answers will be entered as missing. An example of an invalid answer is when a person gives two answers that are incompatible.

Example of an Invalid Answer

3. I believe it should be equally unacceptable for women and men to have sex with casual acquaintances.	Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Mostly Agree	Strongly Agree

- Any **handwritten** answers that do not directly correspond to one of the choices will be entered as missing. Do <u>not</u> try to interpret their answer. For example, a student might write an answer in the margin.

Example of a Handwritten Answer

It depends on how long they've been dating.	4. I believe a woman should be careful not to appear smarter than the man she is dating.	Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Mostly Agree	Strongly Agree

TASK 2: PREPARE YOUR SURVEY/DATA FORMS

After you create your data dictionary, you are ready to prepare the actual surveys for data entry. But before turning to this task, an important reminder about pre-post-follow-up analyses:

- You must be able to match their pre-, post- and follow-up surveys to track how people change in their knowledge, attitudes and/or behaviors over time.
- So you will have all of the pre-survey, post-survey and follow-up survey answers in the SAME spreadsheet.
- Each person = one row. You should be able to read across the row to see the person's answers on all of the surveys.
- For example, if Maria's answers are on Row 4, you will see all of her pre-survey answers there, further to the right you will see all of her post-survey answers, and then furthest to the right will be all of her follow-up survey answers. ALL of her answers will be in Row 4 for ALL of the surveys.

This requires that when you collect the data you have each person write an **identifying number** on their surveys so you can **match them up**. The number should:

- Be easy to remember.
- Be unique to each person so you don't have more than one person with the same number.
- Not change over time.
- Maintain anonymity as best as possible.

Meeting all of these requirements is more challenging than it might seem. The easiest option is to have participants use the last four digits of their home or cell phone number. In theory, if you worked hard enough you might be able to figure out who a survey belonged to, but it would take extraordinary effort. You can also have people create a unique number by adding together numbers such as their birthdate and phone number. However, even simple addition instructions often result in errors and unmatched surveys.

Assuming that you are collecting at least pre- and post-surveys and that you had participants write identifying numbers on them, you are ready to prepare your surveys for data entry.

A. Match the Pre- and Post-Surveys

Go through the surveys and **use the identifying numbers to match up** each person's pre- and post-surveys. It takes a little time to do this, but is worth the time you will save in the end. Matching the surveys now will:

- Make the data entry go faster.
- Let you set aside unmatched surveys so you don't enter those surveys into your spreadsheet.

B. Add Identifying Numbers to Single Surveys

If you are <u>not</u> doing pre-post analyses but instead are analyzing data from a single survey (e.g., a program satisfaction survey), write your own identifying numbers on each survey. This is done by numbering the surveys sequentially. Identifying numbers will help you find surveys if you need to go back and correct any data entry errors.

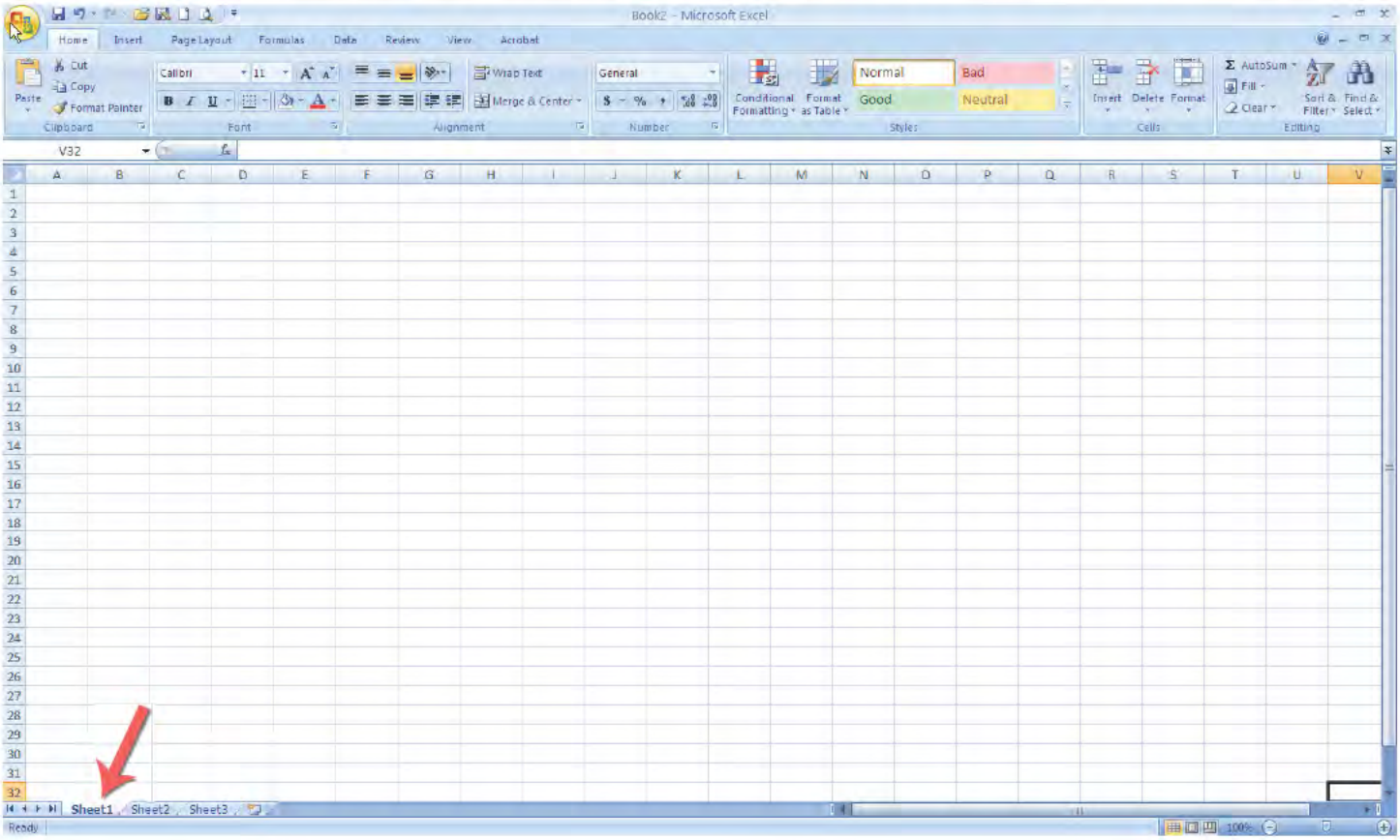
TASK 3: SET UP YOUR WORKSHEET

You are almost ready to enter your data into a spreadsheet, but first you need to set up the spreadsheet. A Microsoft Excel file is called a **workbook**. Within the workbook you can have multiple pages called **worksheets**. Think of these as separate sheets of paper that are stapled together.

A. Orient Yourself to the Worksheets

The default in Excel is to start with three sheets that are labeled "Sheet 1," "Sheet 2," and "Sheet 3."

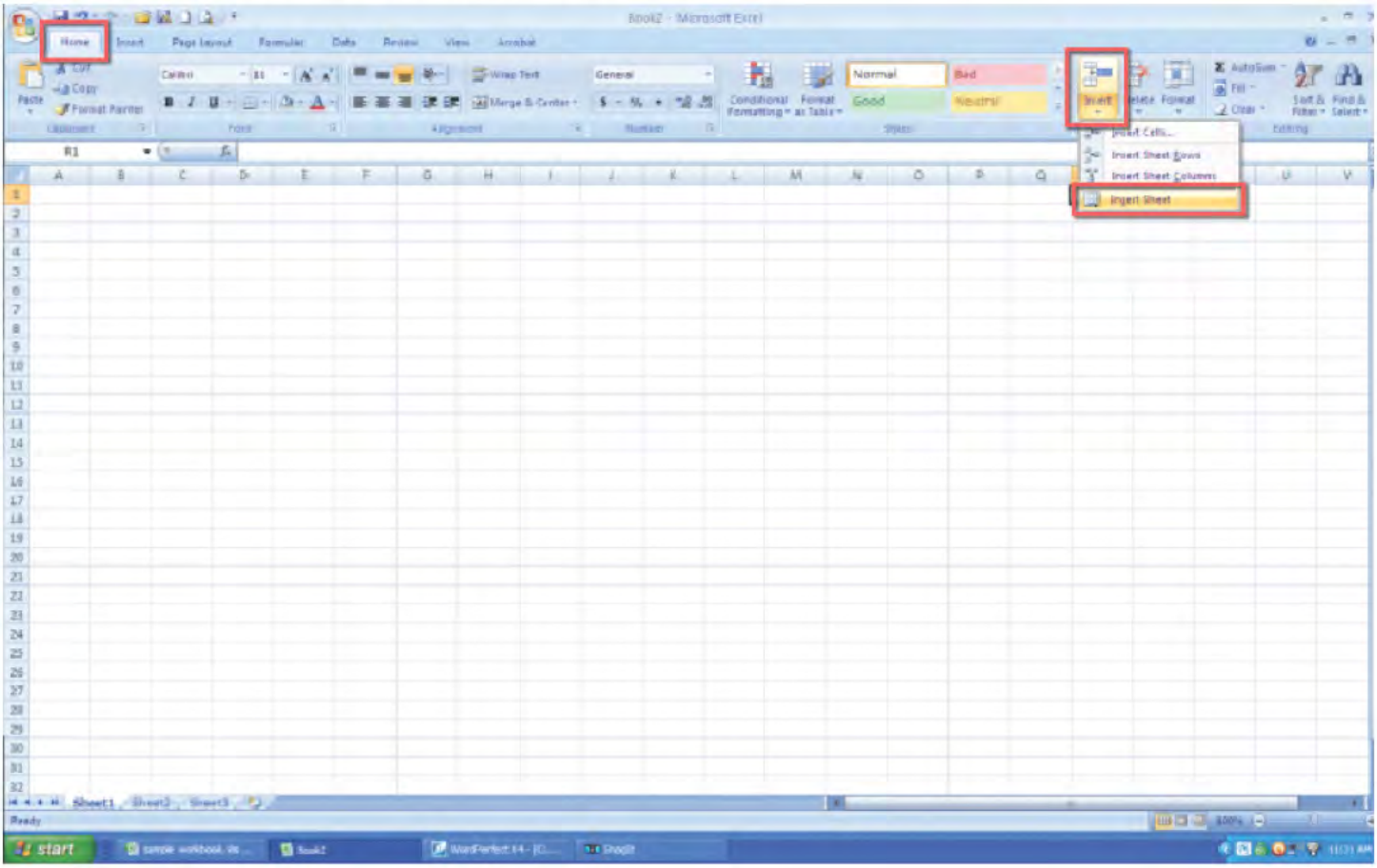
- You can switch from one sheet to another by clicking on the sheet tab at the bottom of the screen.



The default in Excel is to give you three sheets to start with, but you can add or delete as many sheets as you want.

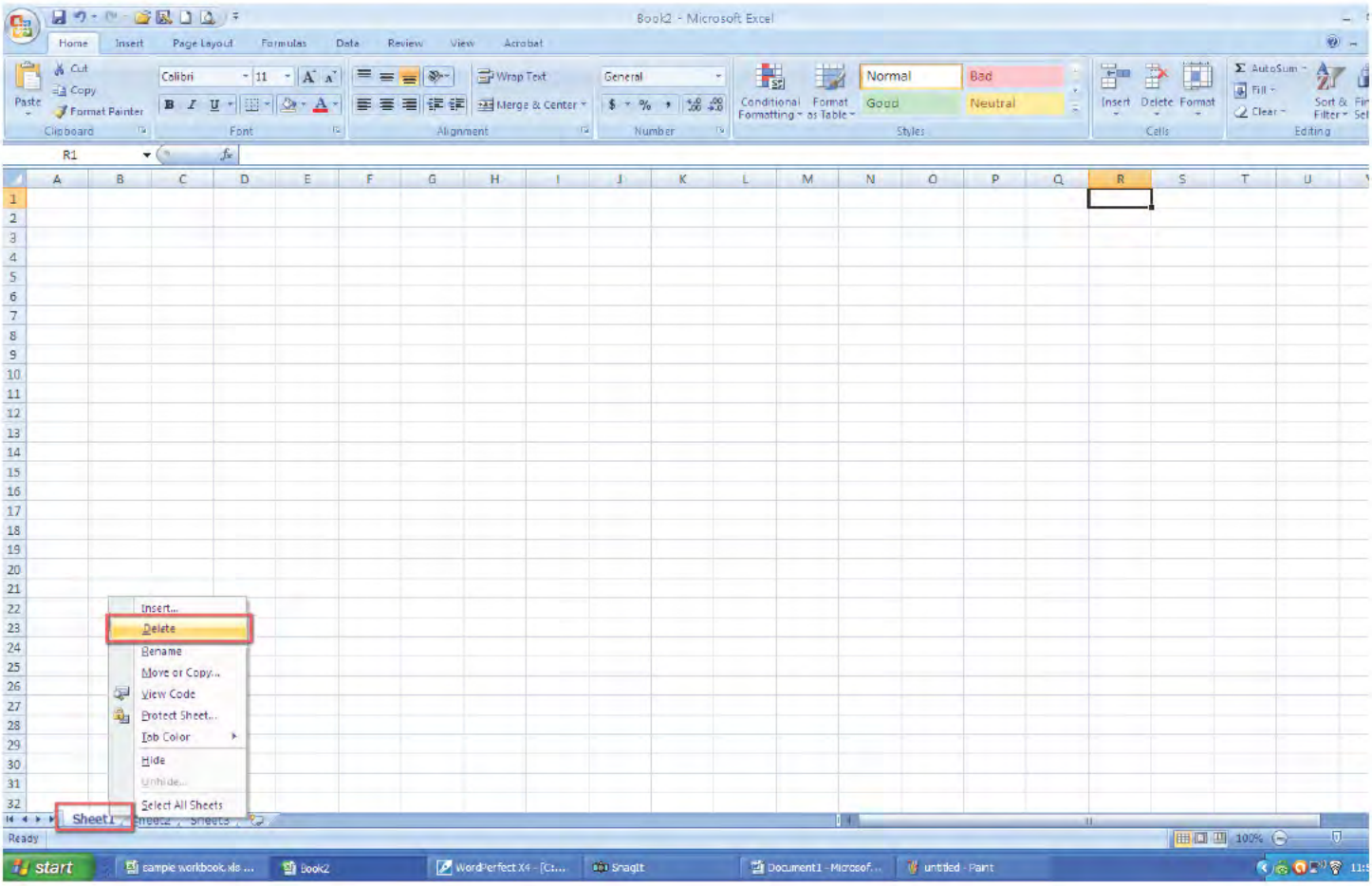
- To add a worksheet:
  - Click on the Home Ribbon (if you're not already there).
  - Click on the Insert button.
  - Click on Insert Sheet.

  The new worksheet will automatically appear (look at the sheet tabs at the bottom of the screen) and will be given the next number in the sequence.



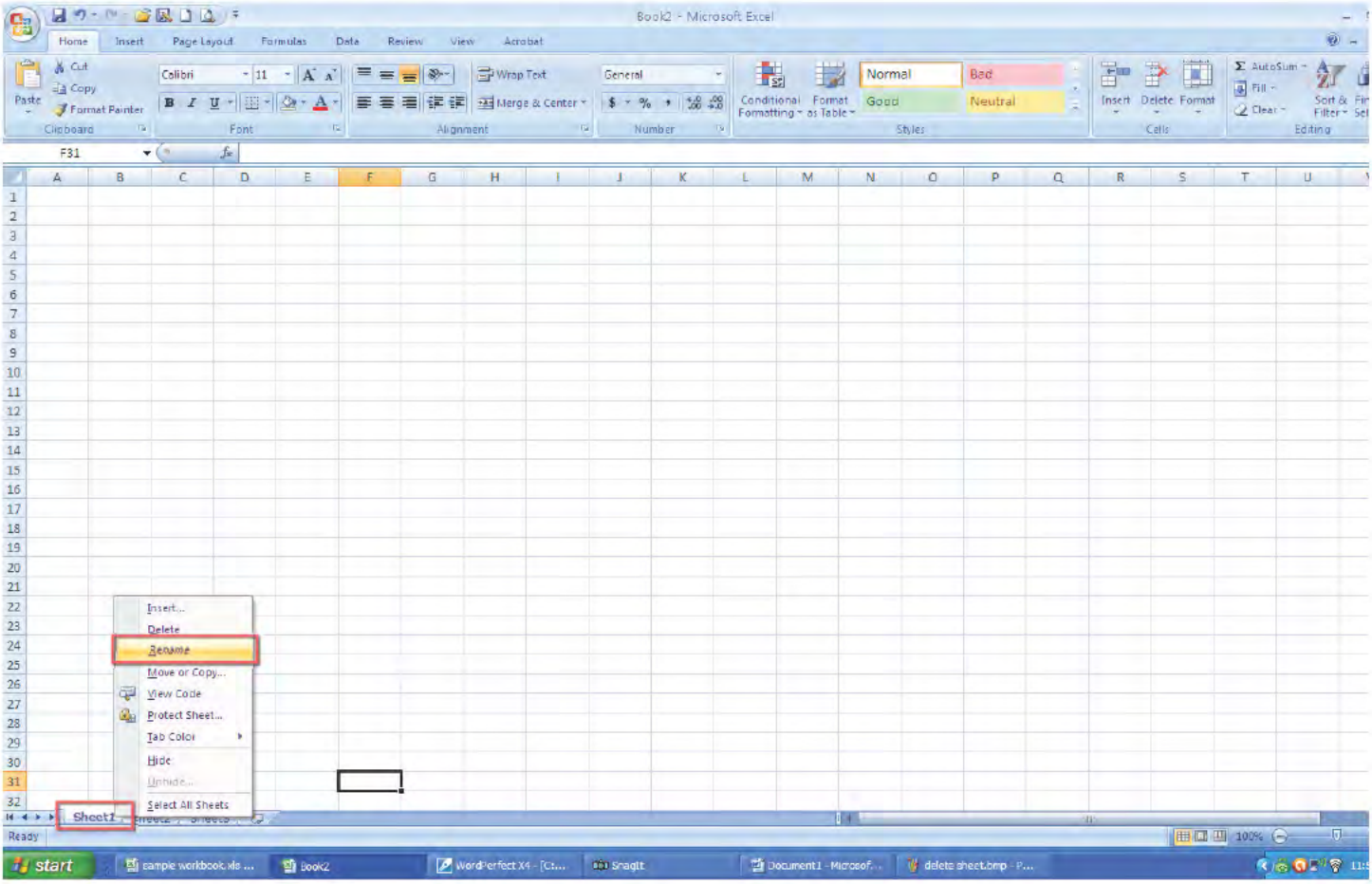
- To delete a worksheet:
  - Right-click on the sheet tab that you want to delete.
  - Click on Delete.
  - The worksheet will automatically disappear.

Note: Only delete a worksheet if you are willing to lose any data or figures on that worksheet.



- To rename a sheet:
  - Right-click on the sheet tab that you want to rename.
  - Click on Rename. This will highlight the current name of the worksheet.
  - Type in the name you want to give it.

Recommendation: Rename Sheet 1 to "Survey Data" or some other appropriate name. This will make it easier later when we want to analyze the data. To keep the worksheets organized we will put the data analysis on a separate worksheet.

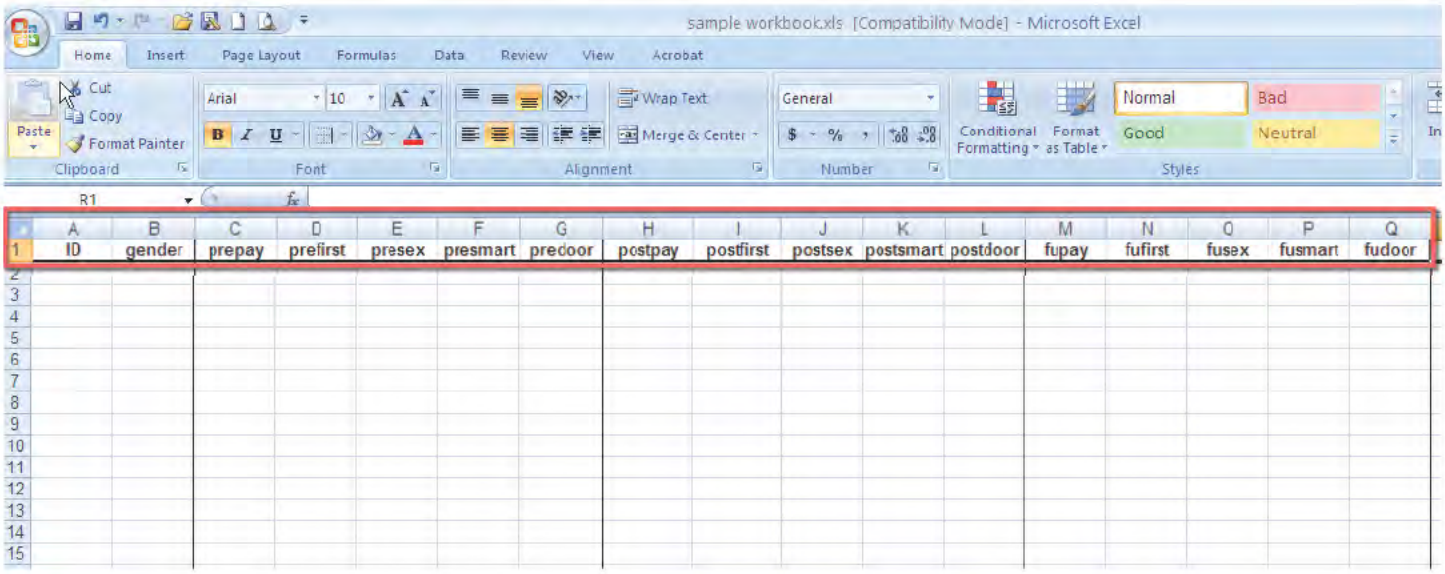


B. Label Your Columns

You are now ready to **label your columns with the labels you wrote on your Data Dictionary.** To do this:

- Type the labels on the top row of the spreadsheet.
- Make sure the ID number is the first column.
- Enter the question labels in the same order as they appear on the surveys. This will make data entry go faster and be more accurate.
- Enter all of the labels for your pre-survey questions. Then enter the labels for your post-survey questions. Finally, enter the labels for your follow-up questions.
- If you collected demographic data (e.g., age, class, gender, etc.), those columns only need to be entered once.

For example, the CRCC spreadsheet would look like this:



There are a few things you can do to make the spreadsheet easier to read:

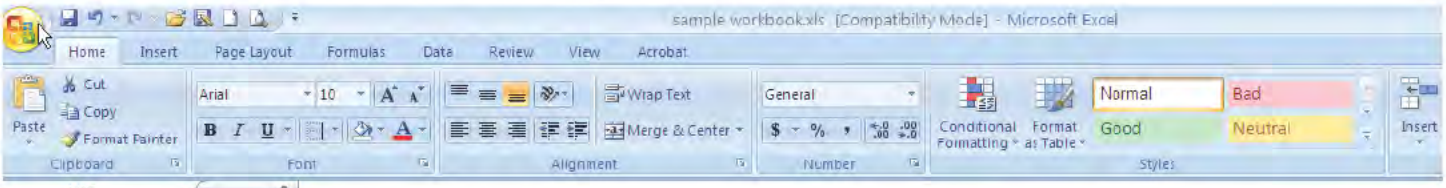
- **Bold the labels** so they stand out from the data you will enter in the next step.
  - *To do this:* Click on the row number (not the ID number, but the row number that Excel automatically puts on the far left side of the screen).
    - This will highlight the entire row.
    - Click on the "B" button on your Home ribbon.
- **Add a dividing line** between the labels in the first row and where you will start to enter data in the second row.
  - *To do this:* Click on the row number (not the ID number) that Excel automatically puts on the far left side of the screen.
    - This will highlight the entire row.
    - Add a border by clicking on the Border button on your Home ribbon and selecting where you want the line to go.
- **Add dividing lines** between the demographic data and other data, between the last pre-survey and the first post-survey columns, and between the last post-survey and first follow-up columns.
  - *To do this:* Click on the column letter (that Excel automatically puts on the top of the worksheet) for the last column for the section.
    - This will highlight the entire column.
    - Add a border by clicking on the Border button on your Home ribbon and selecting where you want the line to go.

TASK 4: ENTER YOUR DATA

Congratulations! You are now ready to enter your data. This step is easy. It simply requires taking each survey and entering each person's data into the corresponding columns. Remember to:

- **Keep each person's pre-, post- and follow-up surveys in the same row.** If Jackie is ID #1234 then when we read across the row for ID #1234 we should see all of Jackie's answers for all of the surveys.
- **Put a check mark on the survey after the data is entered.** This will help you keep track of where you left off and help verify that all surveys have been entered.
- **Proceed slowly.** Rushing through data entry is likely to result in more errors. That will take more time in the next step when you clean your data.
- **Save early and often.** You never know when there will be a temporary power outage, your computer will freeze up or some other technological disaster will occur. There is nothing more frustrating than entering a large amount of data and then having to enter it all again because you lost it. Clicking on the Save button or pressing CTRL+S only takes a few seconds. Make a habit of saving your data after every few surveys. This way you will never lose more than a few minutes of work.
- **Watch out for fatigue.** Trying to enter too much data at one time will increase errors. Make sure you are comfortable and that you take frequent breaks to stretch and rest your eyes.
- **Use staff resources wisely.** Data entry is a great job for a volunteer who is attentive to details and who is willing to ask questions when they arise.

For example, CRCC's completed spreadsheet would look like this:



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	ID	gender	prepay	prefirst	presex	presmart	predoor	postpay	postfirst	postsex	postsmart	postdoor	fupay	fufirst	fusex	fusmart	fudoor	
2	1234	1	2	3	1	3	1	3	4	3	4	4	3	4	3	4	4	
3	2345	2	1	2	1	1	1	3	4	4	4	4	3	4	4	4	4	
4	3456	2	2	3	2	2	1	3	4	3	4	4	3	4	3	4	4	
5	4567	1	4	5	4	4	4	4	5	4	5	4	4	5	4	5	4	
6	5678	2	3	2	4	3	2	4	5	4	4	5	4	5	4	4	5	
7	6789	2	3	3	3	3	3	5	5	5	5	5	5	5	5	5	5	
8	7891	1	2	3	1	3	1	3	4	3	4	4	1	2	1	1	1	
9	8912	2	1	2	1	1	1	3	4	4	4	4	2	3	2	2	1	
10	9123	1	2	3	2	2	1	3	4	3	4	4	4	5	4	4	4	
11	1472	1	4	5	4	4	4	4	5	4	5	4	3	2	4	3	2	
12	4725	2	3	2	4	3	2	4	5	4	4	5	3	3	3	3	3	
13	7258	1	3	3	3	3	3	5	5	5	5	5	3	4	3	4	4	
14	2583	2	2	3	1	3	1	3	4	3	4	4	3	4	4	4	4	
15	5836	1	1	2	1	1	1	3	4	4	4	4	3	4	3	4	4	
16	8369	1	2	3	2	2	1	3	4	3	4	4	4	5	4	5	4	
17	3691	2	4	5	4	4	4	4	5	4	5	4	4	5	4	4	5	
18	6914	2	3	2	4	3	2	4	5	4	4	5	5	5	5	5	5	
19	9147	1	3	3	3	3	3	5	5	5	5	5	5	5	5	5	5	
20	9876	2	3	3	3	3	3	5	5	5	5	5	5	5	5	5	5	
21	8745	2	3	3	3	3	3	5	5	5	5	5	5	5	5	5	5	
22																		
23																		

TASK 5: CLEAN YOUR DATA

No matter how careful you were when you entered your data, it is very common for mistakes to occur. It is important to minimize these errors because they can radically change the results you will get when you run your analyses.

The most common (and serious) mistake is to enter invalid data. This usually happens when you fail to move to the next cell while entering data. For example, if the real answers on the survey were:

prepay = 3

prefirst = 2

You might accidentally enter:

prepay = 32

It can also happen when you actually enter the same number two times. For example, entering:

prepay = 33

instead of prepay = 3

Usually you will catch these errors during the data entry process itself because you will get to the end of the section or survey and realize that you are in the wrong column. (This is why adding dividing lines between sections or surveys is helpful.) However, sometimes you will not catch the error or you may simply hit the wrong key, such as entering "7" when you meant to enter "6." These invalid (or out-of-range) errors must be corrected.

Option A: Scan

At the very least, visually scan your spreadsheet to look for any invalid values. This is an effective strategy if you have a small amount of data.

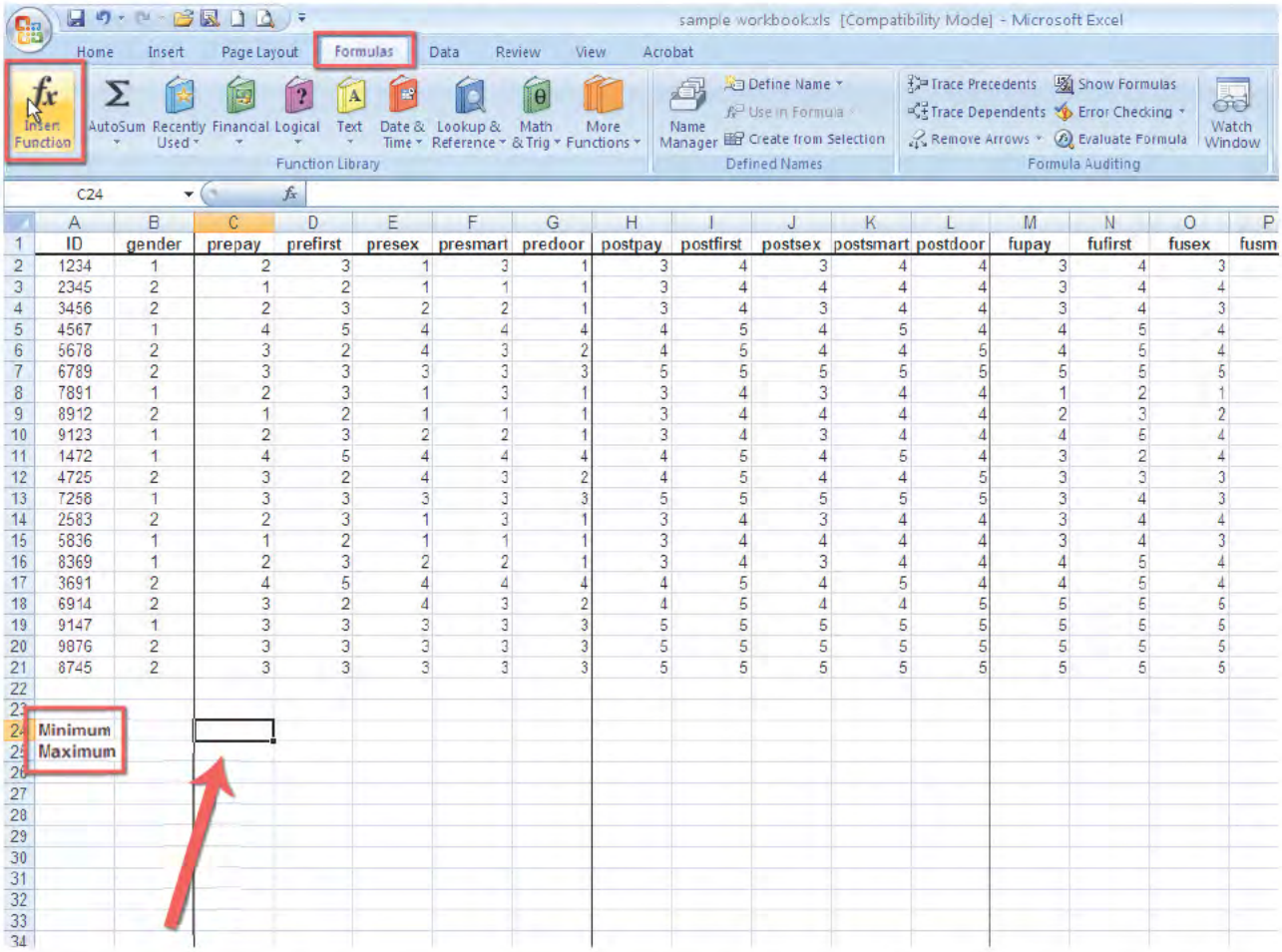
- If you find errors, use the ID number to **find the original survey**. Confirm what the correct answer is and make the correction on the spreadsheet.
- Do **not assume** that you know what the correction should be. A common saying among researchers is: "Garbage in. Garbage out." In other words, your results are only as good as the data you put into the worksheet. It is worth the few minutes it takes to verify what the actual answer was on the survey before making any correction.

Option B: Calculate Minimum and Maximum Values

If you have a large amount of data you probably do not want to rely on a visual scan because it will be too easy to miss the invalid data. In this case, a simple way to catch out-of-range values is to **calculate the minimum and maximum values for each column.** This will tell you whether you entered any out of range values.

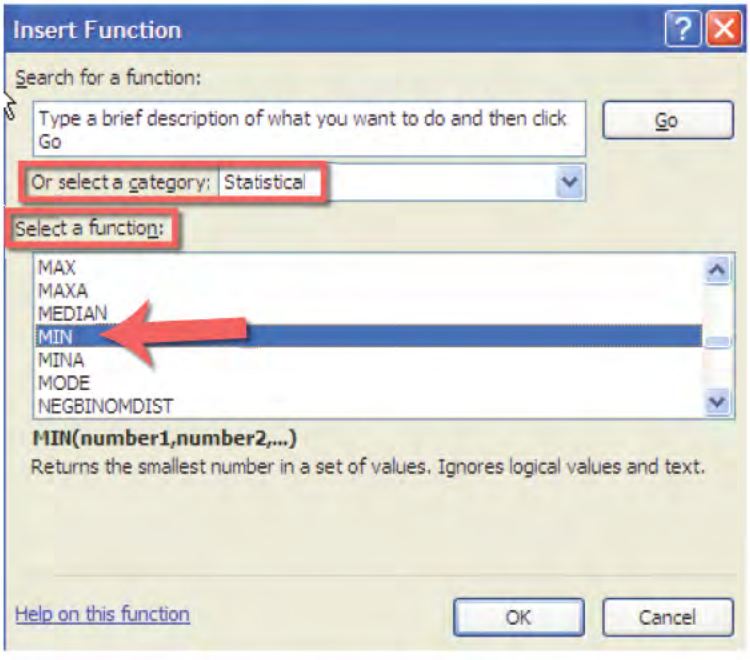
- To calculate minimum and maximum values, do the following:
  - A couple of rows after the last person's data, label one row as "Minimum" and one row as "Maximum."
  - Click on the cell on the "Minimum" row and in the first column of data. This is where the lowest value entered in that column will go.
  - Click on the Formula ribbon and then click on the Insert Function (fx) button. This will open up the Formula dialog box (see next page.)
  - **Click on the cell where the value will go before you click Insert Function.**
  - **Add "Minimum" and "Maximum" labels.**
  - **Click on the Formulas ribbon.**

- Click on the Insert Function button



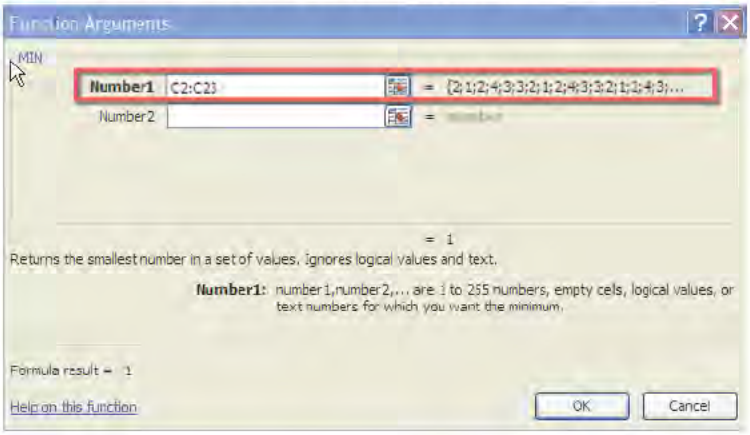
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	ID	gender	prepay	prefirst	presex	presmart	predoor	postpay	postfirst	postsex	postsmart	postdoor	fupay	fufirst	fusex	fusm
2	1234	1	2	3	1	3	1	3	4	3	4	4	3	4	3	
3	2345	2	1	2	1	1	1	3	4	4	4	4	3	4	4	
4	3456	2	2	3	2	2	1	3	4	3	4	4	3	4	3	
5	4567	1	4	5	4	4	4	4	5	4	5	4	4	5	4	
6	5678	2	3	2	4	3	2	4	5	4	4	5	4	5	4	
7	6789	2	3	3	3	3	3	5	5	5	5	5	5	5	5	
8	7891	1	2	3	1	3	1	3	4	3	4	4	1	2	1	
9	8912	2	1	2	1	1	1	3	4	4	4	4	2	3	2	
10	9123	1	2	3	2	2	1	3	4	3	4	4	4	5	4	
11	1472	1	4	5	4	4	4	4	5	4	5	4	3	2	4	
12	4725	2	3	2	4	3	2	4	5	4	4	5	3	3	3	
13	7258	1	3	3	3	3	3	5	5	5	5	5	3	4	3	
14	2583	2	2	3	1	3	1	3	4	3	4	4	3	4	4	
15	5836	1	1	2	1	1	1	3	4	4	4	4	3	4	3	
16	8369	1	2	3	2	2	1	3	4	3	4	4	4	5	4	
17	3691	2	4	5	4	4	4	4	5	4	5	4	4	5	4	
18	6914	2	3	2	4	3	2	4	5	4	4	5	5	5	5	
19	9147	1	3	3	3	3	3	5	5	5	5	5	5	5	5	
20	9876	2	3	3	3	3	3	5	5	5	5	5	5	5	5	
21	8745	2	3	3	3	3	3	5	5	5	5	5	5	5	5	
22																
23																
24	Minimum															
25	Maximum															
26																
27																
28																
29																
30																
31																
32																
33																
34																

- In the Function Dialog Box, where it says "Or select a category," click on the arrow and use the drop-down list to select "Statistical."
- In the "Select a function" box, scroll down until you find the MIN function. Highlight MIN and then click OK.



After you click OK, a dialog box will open up that tells you what cells the function will be performed on. In other words, this is where in the spreadsheet the computer will look to find the minimum value.

The cells are referred to by their column letter and row number. The colon in between the cells means that Excel will look at all of the cells in between those two cells. For example, in the picture below in the "Number 1" box it says **C2:C22**. This means Excel is about to look for the minimum value starting with cell C2 and going through cell C22.[3]



- **Check** that the cell range is correct. *Note:* Excel will do its best to guess what range you want. In this example, the range for the first function you enter will probably be guessed correctly because there are a bunch of numbers, then a blank cell, and then the cell where the answer is going to go. However, in this example when you go to calculate the maximum score below it, Excel would probably guess that you only want the maximum score for C23. This is <u>not</u> what you would want so you would need to change that range to C2:C22.
- If you need to **change the cell range** you can do so either by typing it in the **Number 1** box **OR** by using the mouse to highlight the cells you want the computer to consider.
- After you have checked (and corrected, if needed) the cell range you want Excel to look in, **click OK**. You will see that the minimum value for that column now appears on the "Minimum" row for that column.

- Once you have calculated the minimum score for the first column of data, you can automatically repeat that function for <u>all</u> of the data columns. You can do this by holding your mouse so the cursor hovers over the <u>lower right</u> corner of that cell where you have a minimum score. You will see that a + sign appears. Hold down on the mouse and drag that + sign across the row, highlighting all of the cells as you go. When you let go, the minimum value for all of the columns will appear in the appropriate places.

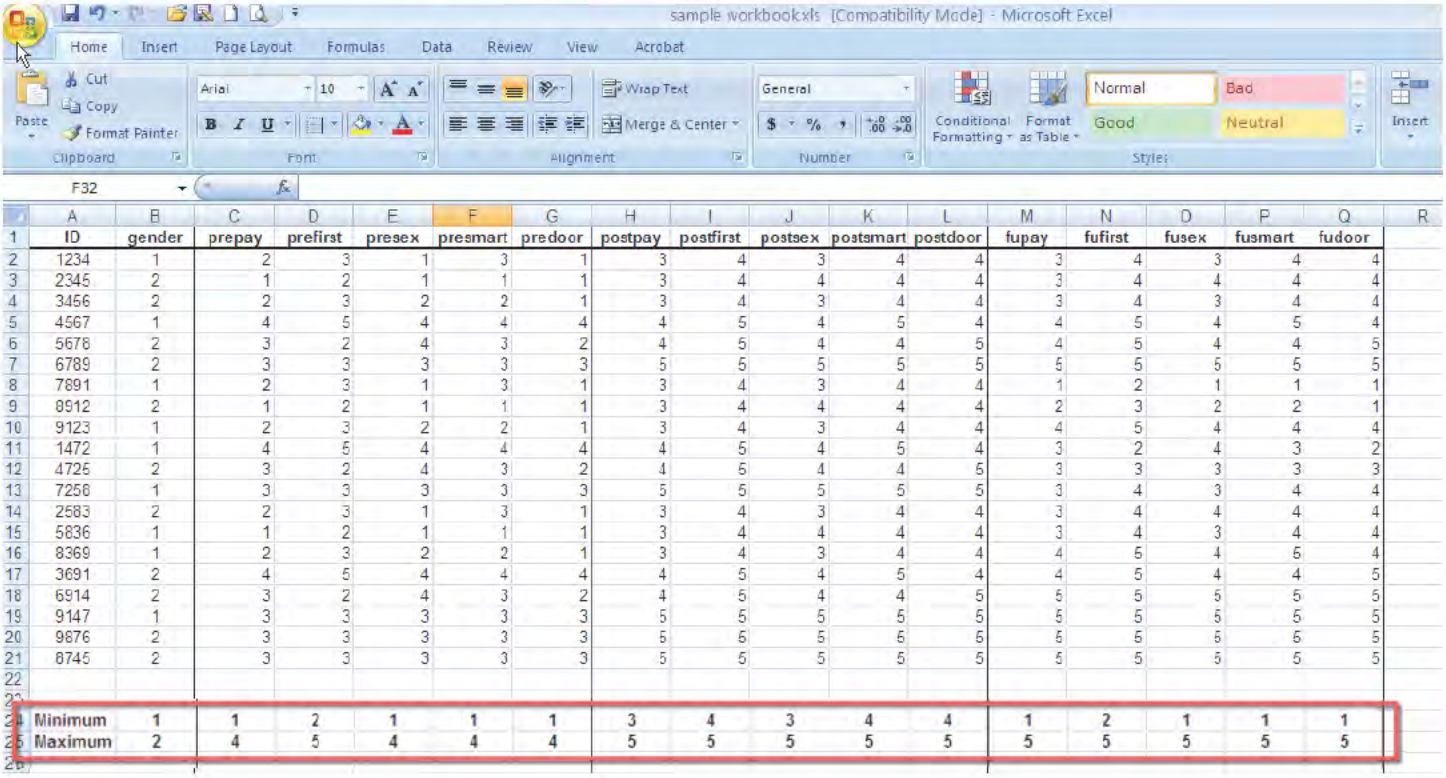
- Repeat these steps on the next row to calculate the maximum score by selecting the MAX function in the Function dialog box.

---

3 Throughout this manual instructions will be given using the point-and-click method and dialog boxes. Those who are more comfortable with computer codes may prefer to type functions in directly. Functions can be typed in the data entry box that appears immediately above the worksheet. For example, in this picture you will see that the function is written as: =MIN(C2:C22). Functions always begin with an equal sign followed by the function command. The cell range follows in parentheses. Sometimes additional criteria information will be required.

- Finally, look at the minimum and maximum scores for all of the columns and verify that there is no number lower or higher than the valid range for each question.

For example, CRCC's spreadsheet would now look like this:



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	ID	gender	prepay	prefirst	presex	presmart	predoor	postpay	postfirst	postsex	postsmart	postdoor	fupay	fufirst	fusex	fusmart	fudoor	
2	1234	1	2	3	1	3	1	3	4	3	4	4	3	4	3	4	4	
3	2345	2	1	2	1	1	1	3	4	4	4	4	3	4	4	4	4	
4	3456	2	2	3	2	2	1	3	4	3	4	4	3	4	3	4	4	
5	4567	1	4	5	4	4	4	4	5	4	5	4	4	5	4	5	4	
6	5678	2	3	2	4	3	2	4	5	4	4	5	4	5	4	4	5	
7	6789	2	3	3	3	3	3	5	5	5	5	5	5	5	5	5	5	
8	7891	1	2	3	1	3	1	3	4	3	4	4	1	2	1	1	1	
9	8912	2	1	2	1	1	1	3	4	4	4	4	2	3	2	2	1	
10	9123	1	2	3	2	2	1	3	4	3	4	4	4	5	4	4	4	
11	1472	1	4	5	4	4	4	4	5	4	5	4	3	2	4	3	2	
12	4725	2	3	2	4	3	2	4	5	4	4	5	3	3	3	3	3	
13	7258	1	3	3	3	3	3	5	5	5	5	5	3	4	3	4	4	
14	2583	2	2	3	1	3	1	3	4	3	4	4	3	4	4	4	4	
15	5836	1	1	2	1	1	1	3	4	4	4	4	3	4	3	4	4	
16	8369	1	2	3	2	2	1	3	4	3	4	4	4	5	4	5	4	
17	3691	2	4	5	4	4	4	4	5	4	5	4	4	5	4	4	5	
18	6914	2	3	2	4	3	2	4	5	4	4	5	5	5	5	5	5	
19	9147	1	3	3	3	3	3	5	5	5	5	5	5	5	5	5	5	
20	9876	2	3	3	3	3	3	5	5	5	5	5	5	5	5	5	5	
21	8745	2	3	3	3	3	3	5	5	5	5	5	5	5	5	5	5	
22																		
23																		
24	**Minimum**	**1**	**1**	**2**	**1**	**1**	**1**	**3**	**4**	**3**	**4**	**4**	**1**	**2**	**1**	**1**	**1**	
25	**Maximum**	**2**	**4**	**5**	**4**	**4**	**4**	**5**	**5**	**5**	**5**	**5**	**5**	**5**	**5**	**5**	**5**	
26																		

- **Scan** the minimum and maximum values to verify that there are no numbers that are **out of range**.

  In the CRCC example, gender was coded as 1 or 2 and when we look at the minimum and maximum values for the gender column we see that all values are within that range. The remaining questions were coded as 1 through 6 and when we look at the minimum and maximum values for those columns we see that all values are within that range. So CRCC staff can move to the next step, knowing that there are no out of range numbers in their worksheet.

  If an out of range number did show up, then we would find where the invalid entry was, look up that person's survey, verify the correct answer and correct the worksheet.

Option C: Preventing Errors with Data Validation

If you are more comfortable with Microsoft Excel, you can use the Data Validation feature to prevent any data entry errors. This feature instructs the computer to accept only values that you specify for each question. If you enter a value that is not valid, you will get an Error Alert message and it will not accept the incorrect data.

The Data Validation feature is activated when you set up your spreadsheet and before you begin entering data.

To activate the Data Validation feature, click on the Data ribbon and select the Data Validation button. This will open up a dialog box where you will use the drop-down menus and blanks to specify what numbers will be accepted.

TASK 6: REVERSE CODE YOUR DATA (IF NECESSARY)

The final task before you begin your data analysis is a process called **reverse coding**. You may not need to do this. However, if you do need to, it is critical. As will be explained below, failure to reverse code answers that need it will result in nonsensical results.

What is reverse coding and why do we do it? To answer this question, let us look at the questions from the CRCC example. For each statement we need to ask ourselves: What does it mean if someone strongly agrees with this statement? In other words, what does it mean if we enter their answer as a "6" in the spreadsheet? The response to this question is indicated to the left of each statement.

"Strongly Agree" means...	DATING AND HETEROSEXUAL RELATIONSHIPS SUBSCALE						
They endorse stereotypical gender roles.	1. I believe a man should be expected to pay the expenses on a date with a woman.	1	2	3	4	5	6
They endorse gender equity.	2. I believe it is all right for a woman to take the first steps to start a relationship with a man.	1	2	3	4	5	6
They endorse gender equity.	3. I believe it should be equally unacceptable for women and men to have sex with casual acquaintances.	1	2	3	4	5	6
They endorse stereotypical gender roles.	4. I believe a woman should be careful not to appear smarter than the man she is dating.	1	2	3	4	5	6
They endorse gender equity.	5. I believe when men show special courtesies only to women (like holding open the door), it reinforces the stereotype that women are helpless.	1	2	3	4	5	6

When we look at these interpretations we see that what the value of "6" for Strongly Agree means in questions 1 and 4 is **opposite** of what it means in questions 2, 3 and 5. But remember our earlier rule about coding: **"Make your codes logically increase in value from low to high where higher numbers indicate "more" of something."** Currently, these codes aren't logical because in one case they mean "more gender equity" and in the other case they mean "more gender stereotypes."

To fix this so that **all questions are interpreted the same way**, we need to **reverse** the codes for questions 1 and 4. The original data we entered for those two questions will be changed so that:

1	will become	6
2	will become	5
3	will become	4
4	will become	3
5	will become	2
6	will become	1

If we replace the worded answers with the **reversed codes** that will be entered in the spreadsheet what we have is:

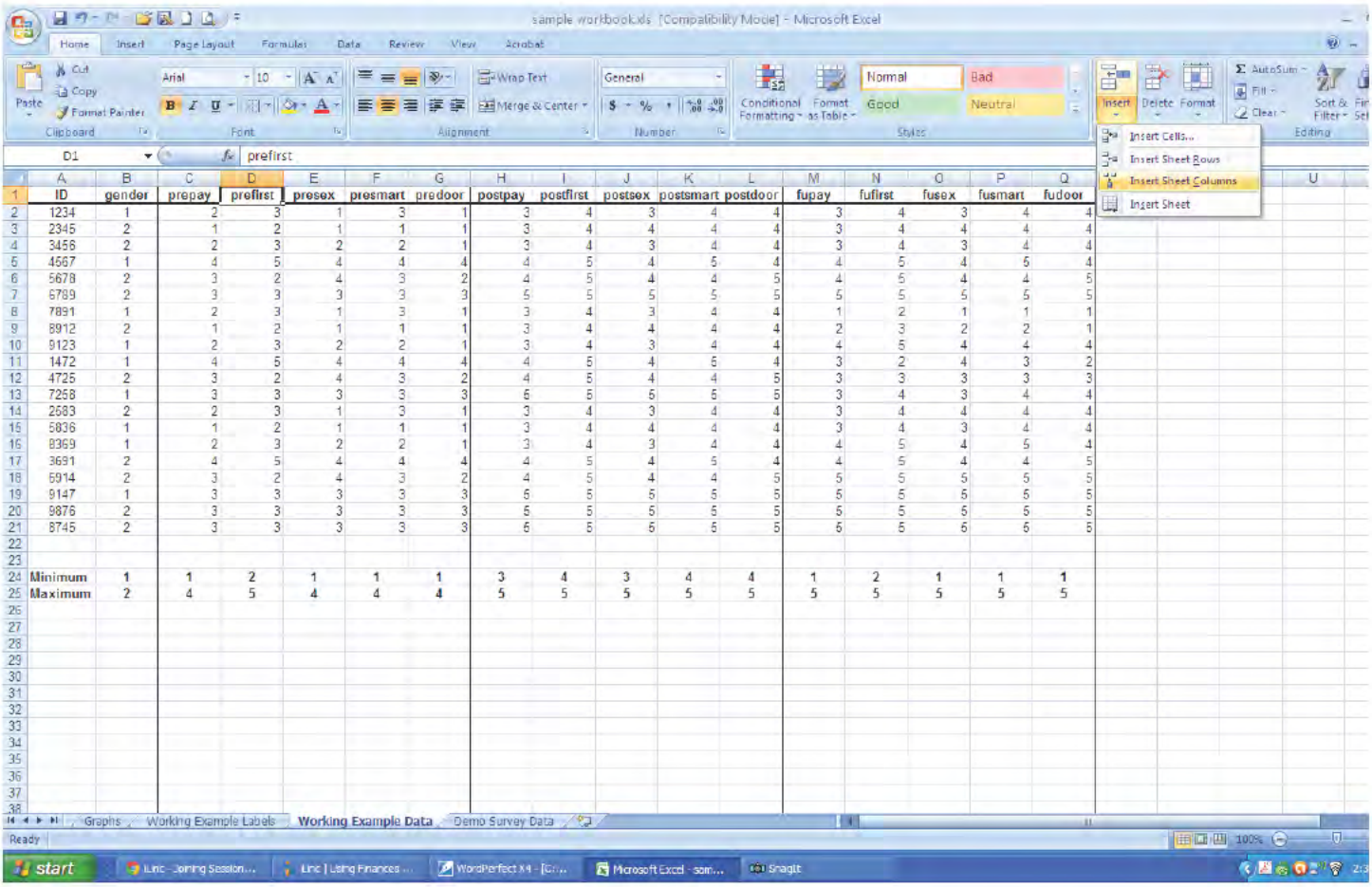
"Strongly Agree" means...	DATING AND HETEROSEXUAL RELATIONSHIPS SUBSCALE						
They endorse gender equity.	1. I believe a man should be expected to pay the expenses on a date with a woman.	6	5	4	3	2	1
They endorse gender equity.	2. I believe it is all right for a woman to take the first steps to start a relationship with a man.	1	2	3	4	5	6
They endorse gender equity.	3. I believe it should be equally unacceptable for women and men to have sex with casual acquaintances.	1	2	3	4	5	6
They endorse gender equity.	4. I believe a woman should be careful not to appear smarter than the man she is dating.	6	5	4	3	2	1
They endorse gender equity.	5. I believe when men show special courtesies only to women (like holding open the door), it reinforces the stereotype that women are helpless.	1	2	3	4	5	6

Because we reverse coded items 1 and 4, we now have a situation where **an answer that is coded as "6" always means the same thing: "6" always means that the person endorses gender equity.** It is possible to reverse code your data **before** entering it into the worksheet. However, this is usually not a good idea. When entering data, your fingers will get into a pattern where you always want to enter the same number for the same position in the table. In other words, you will automatically enter "6" every time you see a mark in the farthest space on the right of the survey.

It is better to reverse code any items that need it after entering your data. To do this, complete the following steps:

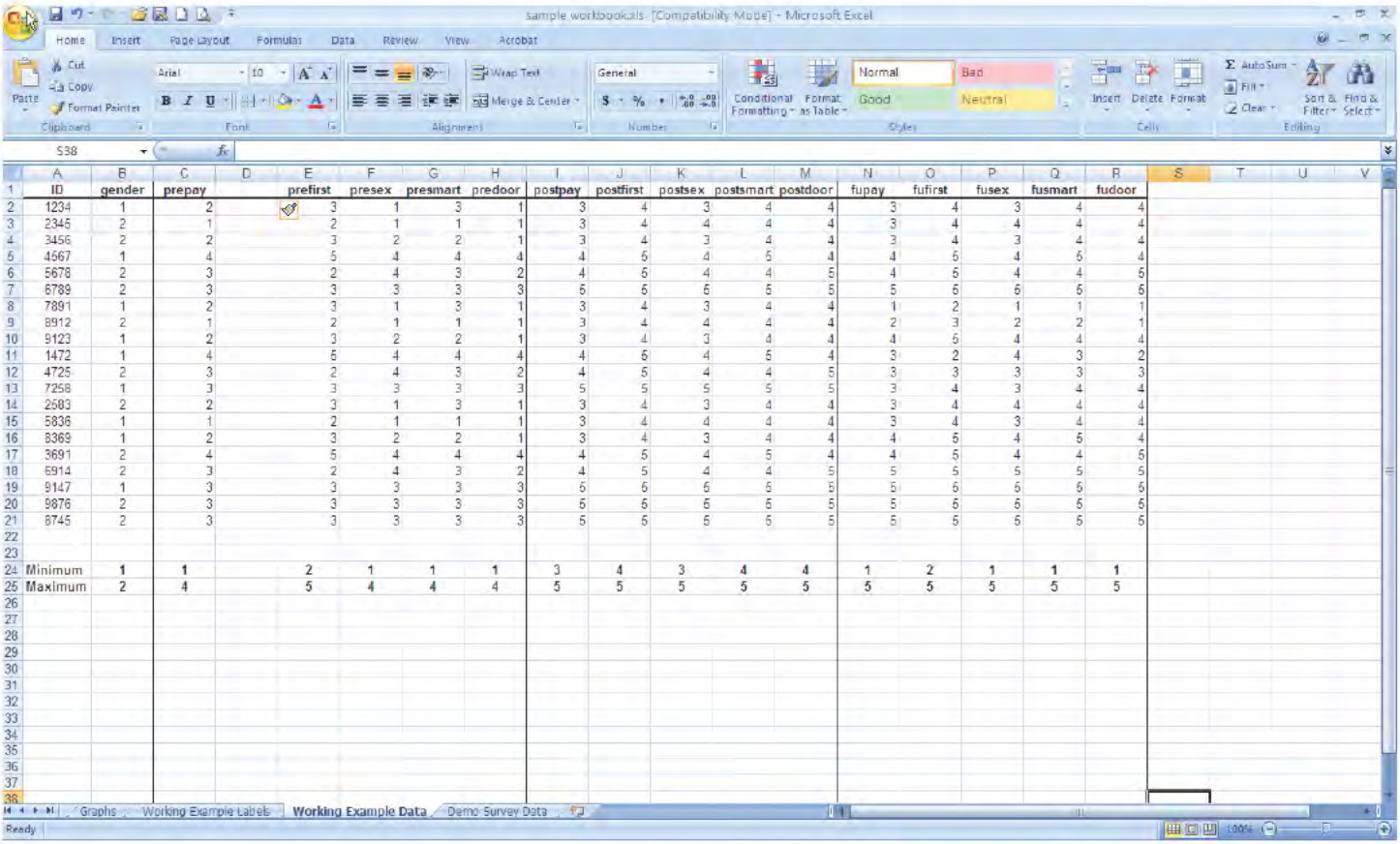
- **Determine which items (if any) need to be reverse coded.** If a higher score is always reflective of the same thing for all of your questions, then no reverse coding is needed and you can skip to the next section.
- Once you have determined which questions need to be reverse coded, **insert a new column in your spreadsheet after each of the original columns for those items** (see picture on next page).

- *To do this:* Click your mouse on the column heading to the right of where you want the new column to go. (For example, to insert a new column after "prepay" you would click on the next column which is "prefirst.")
- Click on the Home ribbon (if you are not already there).
- Click on the Insert button.
- Click Insert Sheet Column.
  - A new column will automatically be inserted after "prepay."



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	ID	gender	prepay	prefirst	presex	presmart	predoor	postpay	postfirst	postsex	postsmart	postdoor	fupay	fufirst	fusex	fusmart	fudoor
2	1234	1	2	3	1	3	1	3	4	3	4	4	3	4	3	4	4
3	2345	2	1	2	1	1	1	3	4	4	4	4	3	4	4	4	4
4	3456	2	2	3	2	2	1	3	4	3	4	4	3	4	3	4	4
5	4567	1	4	5	4	4	4	4	5	4	5	4	4	5	4	5	4
6	5678	2	3	2	4	3	2	4	5	4	4	5	4	5	4	4	5
7	6789	2	3	3	3	3	3	5	5	5	5	5	5	5	5	5	5
8	7891	1	2	3	1	3	1	3	4	3	4	4	1	2	1	1	1
9	8912	2	1	2	1	1	1	3	4	4	4	4	2	3	2	2	1
10	9123	1	2	3	2	2	1	3	4	3	4	4	4	5	4	4	4
11	1472	1	4	5	4	4	4	4	5	4	5	4	3	2	4	3	2
12	4725	2	3	2	4	3	2	4	5	4	4	5	3	3	3	3	3
13	7258	1	3	3	3	3	3	5	5	5	5	5	3	4	3	4	4
14	2683	2	2	3	1	3	1	3	4	3	4	4	3	4	4	4	4
15	5836	1	1	2	1	1	1	3	4	4	4	4	3	4	3	4	4
16	8369	1	2	3	2	2	1	3	4	3	4	4	4	5	4	5	4
17	3691	2	4	5	4	4	4	4	5	4	5	4	4	5	4	4	5
18	5914	2	3	2	4	3	2	4	5	4	4	5	5	5	5	5	5
19	9147	1	3	3	3	3	3	5	5	5	5	5	5	5	5	5	5
20	9876	2	3	3	3	3	3	5	5	5	5	5	5	5	5	5	5
21	8745	2	3	3	3	3	3	5	5	5	5	5	5	5	5	5	5
22																	
23																	
24	Minimum	1	1	2	1	1	1	3	4	3	4	4	1	2	1	1	1
25	Maximum	2	4	5	4	4	4	5	5	5	5	5	5	5	5	5	5

- You will now have a blank column that needs to be filled in with the reverse coded data.
- Label that new column where the reverse coded data will go with the **same label** as the original column **followed by (R)** to show that it contains reverse coded answers. Your new column will look like the following picture:



ID	gender	prepay		prefirst	presex	presmart	predoor	postpay	postfirst	postsex	postsmart	postdoor	fupay	fufirst	fusex	fusmart	fudoor
1234	1	2		3	1	3	1	3	4	3	4	4	3	4	3	4	4
2345	2	1		2	1	1	1	3	4	4	4	4	3	4	4	4	4
3456	2	2		3	2	2	1	3	4	3	4	4	3	4	3	4	4
4567	1	4		5	4	4	4	4	5	4	5	4	4	5	4	5	4
5678	2	3		2	4	3	2	4	5	4	4	5	4	5	4	4	5
6789	2	3		3	3	3	3	5	5	5	5	5	5	5	5	5	5
7891	1	2		3	1	3	1	3	4	3	4	4	1	2	1	1	1
8912	2	1		2	1	1	1	3	4	4	4	4	2	3	2	2	1
9123	1	2		3	2	2	1	3	4	3	4	4	4	5	4	4	4
1472	1	4		5	4	4	4	4	5	4	5	4	3	2	4	3	2
4725	2	3		2	4	3	2	4	5	4	4	5	3	3	3	3	3
7258	1	3		3	3	3	3	5	5	5	5	5	3	4	3	4	4
2583	2	2		3	1	3	1	3	4	3	4	4	3	4	4	4	4
5836	1	1		2	1	1	1	3	4	4	4	4	3	4	3	4	4
8369	1	2		3	2	2	1	3	4	3	4	4	4	5	4	5	4
3691	2	4		5	4	4	4	4	5	4	5	4	4	5	4	4	5
5914	2	3		2	4	3	2	4	5	4	4	5	5	5	5	5	5
9147	1	3		3	3	3	3	5	5	5	5	5	5	5	5	5	5
9876	2	3		3	3	3	3	5	5	5	5	5	5	5	5	5	5
8745	2	3		3	3	3	3	5	5	5	5	5	5	5	5	5	5
Minimum	1	1		2	1	1	1	3	4	3	4	4	1	2	1	1	1
Maximum	2	4		5	4	4	4	5	5	5	5	5	5	5	5	5	5

- You are now ready to **enter the reversed codes** into your new (prepay(R)) column. Have a reference sheet handy to avoid confusion. You might find it easier to reverse all of one number at a time and then go on to the next number, etc.
- **Repeat these steps** for **all items that need to be reverse coded**. Remember to do any necessary reverse coding on all of the surveys (pre-, post- and follow-up surveys).

Note: **Having reverse coded items on surveys is a good practice** because it slows people down, makes them think about their answers and makes it less likely that they will automatically give you the socially desirable answers because it won't be so obvious that the "good" answers are all "Strongly Agree." Once you have practice with reverse coding, it will become easy to think about and to do.

You are now done entering your data and are ready to move on to analyzing it.

CHECKLIST

In this section you learned how to organize and manage your quantitative data from surveys, archival sources and other sources. These tasks were:

Task 1: Create a Data Dictionary

- Label your questions on a blank survey.
- Code the possible answers on a blank survey.
- Decide on and write down rules for coding missing and ambiguous data.

Task 2: Prepare Your Survey/Data Forms

- Match the pre-, post- and follow-up surveys based on unique identifiers.
- Add identifying numbers to any single surveys (e.g., satisfaction surveys).

Task 3: Set Up Your Spreadsheet

- Get familiar with the worksheets.
- Label your columns.
- Add any formatting that will help you read the worksheet.

Task 4: Enter Your Data

- Keep each person's pre-, post- and follow-up surveys in the same row.
- Put a check mark on the survey after the answers are entered.
- Proceed slowly.
- Save early and often.
- Watch out for fatigue.
- Use staff resources wisely.

Task 5: Clean Your Data

- Scan your worksheet for invalid codes.
- Calculate minimum and maximum values and check for out of range values.

Task 6: Reverse Code (if necessary)

- Identify any reverse coding that is needed.
- Insert new columns following the original columns for those items.
- Enter the reverse codes.

- Task 1: Create Scaled Scores
- Task 2: Identify Questions and Type of Data
- Task 3: Identify Appropriate Analysis
- Task 4: Run Analysis

ANALYZING QUANTITATIVE EVALUATION DATA

Data analysis is made easier with the statistical functions provided by programs like Microsoft Excel. When you are first learning the skills it can take a bit of time. But as you gain more experience you may find that this step in the evaluation process is not only easy but also **fun**. If you are someone who likes doing any kind of puzzle, you may be surprised by how much you enjoy data analysis.

Analyzing quantitative data involves four main tasks:

1. Create scaled scores (if necessary).
2. Identify what you want to know and the type of data you have.
3. Identify the appropriate analysis for that data.
4. Run the analysis.

TASK 1: CREATE SCALED SCORES (IF NECESSARY)

Even a short survey or archival measure can quickly lead to an avalanche of data. In the CRCC example we have been working with:

CRCC Survey

5 attitude questions x 3 surveys . . . . . 15 answers per person

Plus gender data . . . . . 16 answers per person

x 20 people . . . . . 320 total answers

Pre-survey compared with post- and follow-up survey questions. . . . . 960 comparisons

Each question on each survey also can be compared to see if there are differences between girls and boys. . . . . 900 comparisons

2180 total points of data!

If we were to try and analyze every question and every possible comparison and to make sense of the sometimes very minor differences between questions, **we would quickly be overwhelmed and unable to make any sense of the results**. Put another way, does it really matter how much attitudes about who pays for the date change, versus how much attitudes about who asks for the data change, versus how much attitudes about appearing smart change, etc.?

What we are really interested in is not the individual questions or individual people but rather the **general attitudes about dating and heterosexual relationships**. Our evaluation question is: Do students come to endorse more gender equity in dating and heterosexual relationships? It's the bigger picture we want to see.

To capture the bigger picture, we need to synthesize (or reduce) the five questions we have asked on each survey into a single score that **summarizes** the students' attitudes about dating and heterosexual relationships. This single score is what we call a **scaled score**.

A **scale** has some particular characteristics:

- It is made of more than one question.
- All the questions measure the same concept.
- All the questions are answered with the same possible responses.
- All the answers are coded the same way.

Dating and Heterosexual Relationships Scale

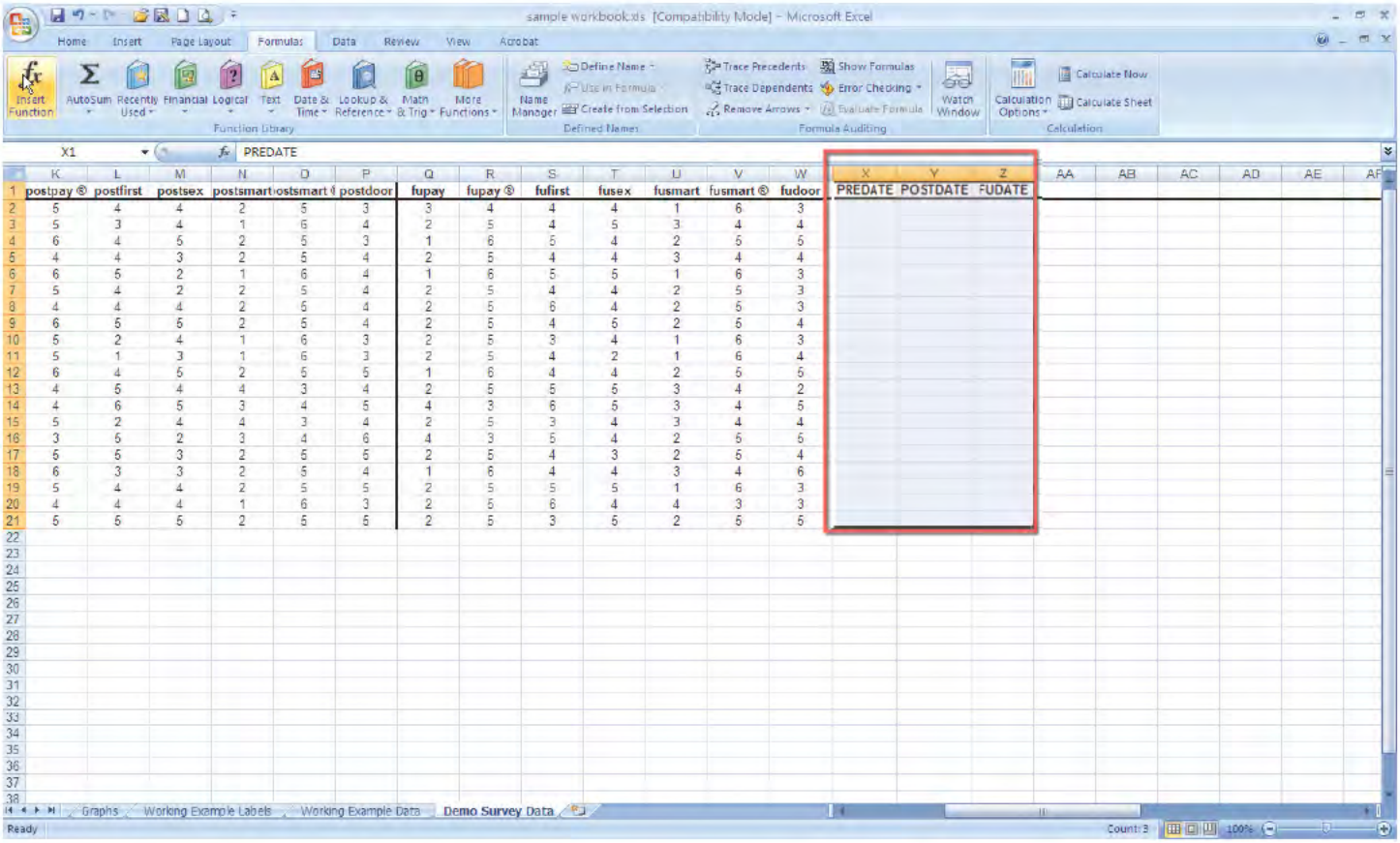
- The scale has five questions.
- All the questions measure attitudes about dating and heterosexual relationships.
- All the questions are answered with the same *Strongly Disagree, Mostly Disagree, Slightly Disagree, Slightly Agree, Mostly Agree, Strongly Agree* responses.
- All the answers are coded from 1–6.
- After reverse coding of questions #1 and #4, higher numbers indicate greater endorsement of gender equity.

To Create a Scaled Score:

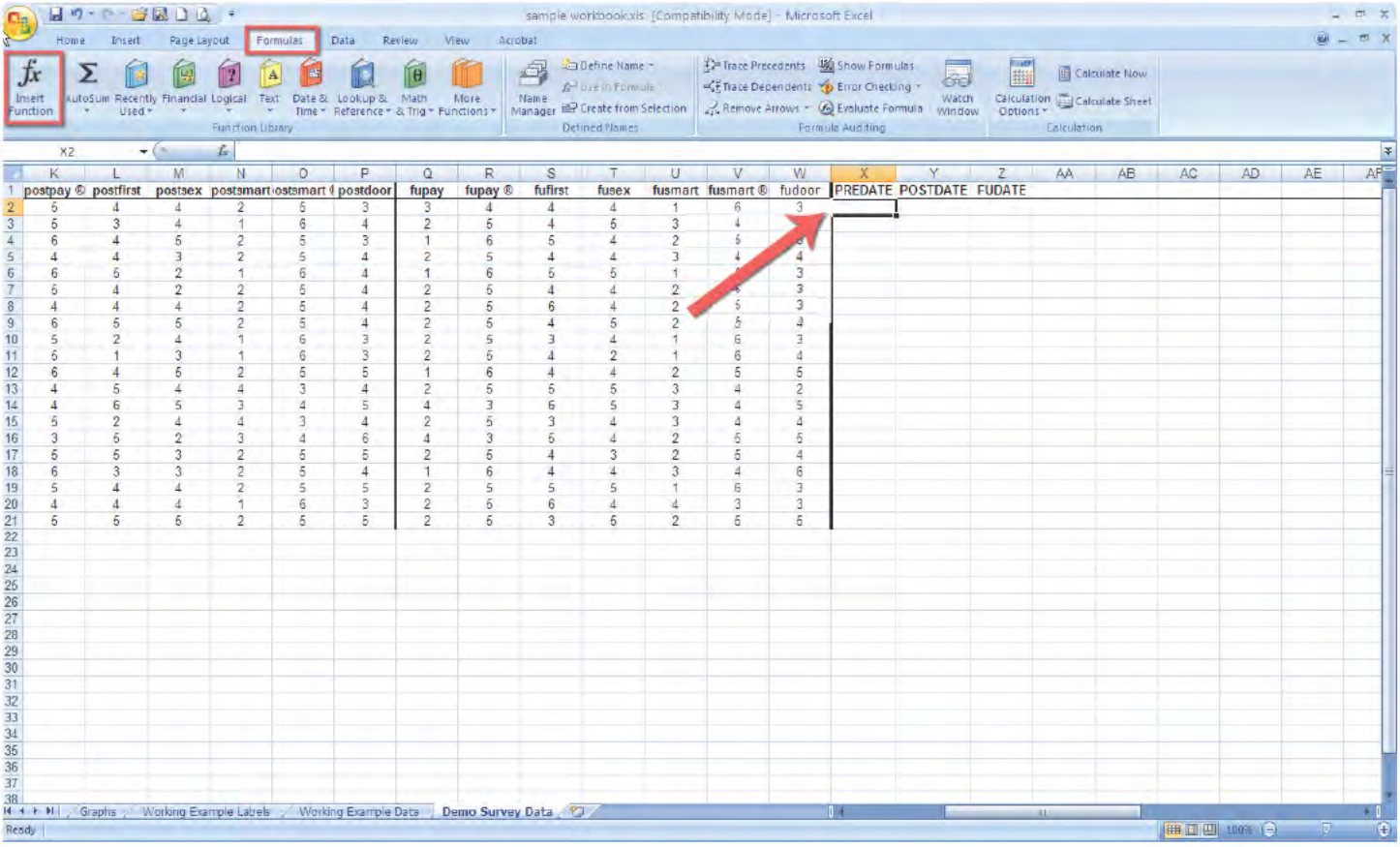
- **Identify which questions make up your scale.** Verify that they meet all of the characteristics described above. The most common mistake is to try and combine questions that measure different concepts.
- **Decide what type of scaled score makes the most sense.** The two most common types are **average (or mean) scores** and **summed scores.**
  - **Average** scores are usually used when you have data like in the CRCC example.
  - **Summed** scores are usually used when you are counting something (e.g., give people a list of 10 strategies for resolving conflicts and ask them to check off how many of them they have used in the past week; then add up the number of positive strategies used).

- **Insert a column in your worksheet for the scaled score and label it.** You can put your new column anywhere in the worksheet that will make it easy for you to find and manage. The most common practice is to put scaled scores at the end of the sheet (i.e., to the far right side). In this case, you do not need to insert a new column; all you need to do is label the first blank column on the right side. Make the label easy to interpret. If you have pre-, post- and follow-up surveys you will need a scaled score for each survey.

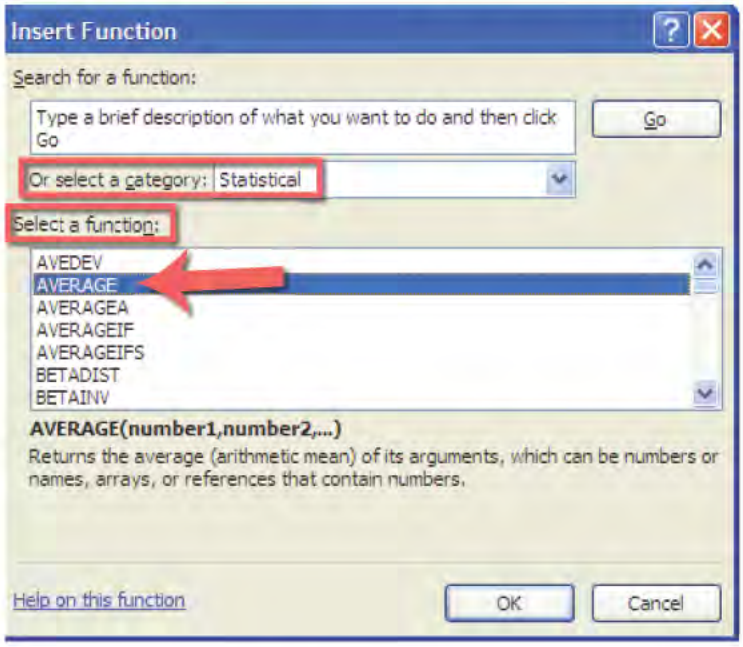
In the CRCC example (shown below) columns were added at the end of the sheet labeled "PREDATE," "POSTDATE" and "FUDATE."



- To fill in those columns with an **average score,** you will use Excel's functions much like you did when you calculated the minimum and maximum scores. However, you will be using the AVERAGE function and you will be defining the range of cells across each row.
  - *To do this:* **Click on the cell** for the first person (row) where you want their average score to go.
  - Click on the **Insert Function button** on the Insert ribbon. This will open up the Function Dialog Box.

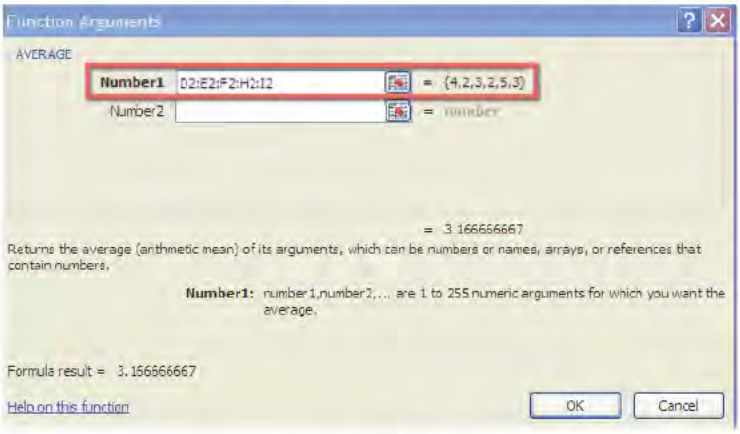


- In the Function Dialog Box make sure that where it says "Or select a category" the category is "Statistical." If it is not, then use the drop-down list to **select the Statistical functions**.
- In the "Select a function" box scroll up until you find the AVERAGE function. **Highlight AVERAGE** and click OK. *Warning:* There are multiple functions for averaging. Make sure you are using the plain AVERAGE function. Do not use the AVEDEV or AVERAGE A. These are different functions and will give you different results.



- After you click OK, a dialog box will open up that tells you what cells the function will be performed on. The cells are referred to by their column letter and row number. The colon in-between the cells means that Excel will look at all of the cells in between those two cells. You can change the cell range by typing in the cell range OR by using your mouse to highlight the cells you want included.

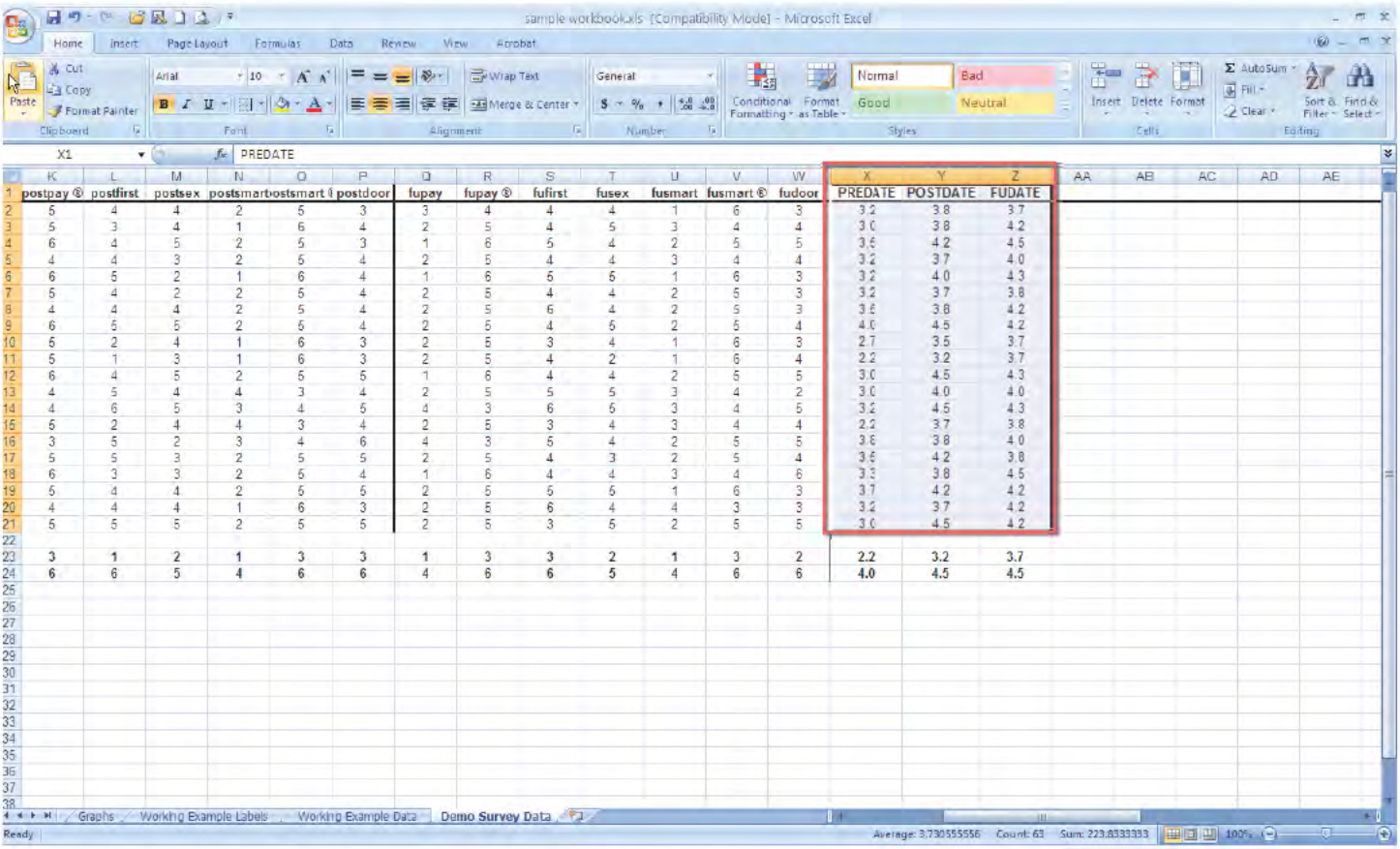
  In the CRCC example, the Pre-test scaled score for the first person will be based on cells D2, E2, F2, H2 and I2. Make sure that if you reverse coded any items you use the reverse coded columns and that you do not use the original columns.



- Click OK. You will see that the average Pre-test score now appears in the PREDATE column for the first person in the worksheet.

- Once you calculate the average score for the first person, you can automatically repeat that function in the column for everyone. You can do this by holding your mouse so the cursor hovers over the lower right corner of that cell where you have the first average score. You will see that a + sign appears. Hold down on the mouse and drag that + sign down the column, highlighting all of the cells as you go. When you let go of the mouse, the average score for each person will appear in the appropriate places.
- Repeat these steps to calculate the average score for all remaining scales on the post and follow-up surveys.

In the CRCC example the scaled scores appear in the picture below.



	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	postpay ®	postfirst	postsex	postsmart	postsmart ®	postdoor	fupay	fupay ®	fufirst	fusex	fusmart	fusmart ®	fudoor	PREDATE	POSTDATE	FUDATE
2	5	4	4	2	5	3	3	4	4	4	1	6	3	3.2	3.8	3.7
3	5	3	4	1	6	4	2	5	4	5	3	4	4	3.0	3.8	4.2
4	6	4	5	2	5	3	1	6	5	4	2	5	5	3.6	4.2	4.5
5	4	4	3	2	5	4	2	5	4	4	3	4	4	3.2	3.7	4.0
6	6	5	2	1	6	4	1	6	5	5	1	6	3	3.2	4.0	4.3
7	5	4	2	2	5	4	2	5	4	4	2	5	3	3.2	3.7	3.8
8	4	4	4	2	5	4	2	5	6	4	2	5	3	3.6	3.8	4.2
9	6	5	5	2	5	4	2	5	4	5	2	5	4	4.0	4.5	4.2
10	5	2	4	1	6	3	2	5	3	4	1	6	3	2.7	3.5	3.7
11	5	1	3	1	6	3	2	5	4	2	1	6	4	2.2	3.2	3.7
12	6	4	5	2	5	5	1	6	4	4	2	5	5	3.0	4.5	4.3
13	4	5	4	4	3	4	2	5	5	5	3	4	2	3.0	4.0	4.0
14	4	6	5	3	4	5	4	3	6	5	3	4	5	3.2	4.5	4.3
15	5	2	4	4	3	4	2	5	3	4	3	4	4	2.2	3.7	3.8
16	3	5	2	3	4	6	4	3	5	4	2	5	5	3.6	3.8	4.0
17	5	5	3	2	5	5	2	5	4	3	2	5	4	3.6	4.2	3.8
18	6	3	3	2	5	4	1	6	4	4	3	4	6	3.3	3.8	4.5
19	5	4	4	2	5	5	2	5	5	5	1	6	3	3.7	4.2	4.2
20	4	4	4	1	6	3	2	5	6	4	4	3	3	3.2	3.7	4.2
21	5	5	5	2	5	5	2	5	3	5	2	5	5	3.0	4.5	4.2
22																
23	3	1	2	1	3	3	1	3	3	2	1	3	2	2.2	3.2	3.7
24	6	6	5	4	6	6	4	6	6	5	4	6	6	4.0	4.5	4.5

- Before moving to the next task you will want to complete the minimum and maximum information for your scaled scores. This information will be used later to describe your results. As you did earlier, use the + and drag across the minimum and maximum rows to ensure that every column's minimum and maximum values are filled in.

TASK 2: IDENTIFY THE TYPES OF DATA

When you need to put new batteries in a flashlight, you first have to know what type of battery the flashlight takes. If you put the wrong type of battery in, it won't fit and the flashlight won't work. Similarly, before you decide what analyses you will run you need to know what type of data you have. Just like you can't fit a size D battery into a size AA flashlight, you can't put data into any statistical test.

There are two main types of data[4]:

- **Categorical (a.k.a. Nominal) data** are those that are categories or names. For example, a person may identify as man, woman or transgender — these are categories of gender. They may identify as Latina or Asian or multiracial — these are categories of race/ethnicity. They may identify as straight, lesbian, gay or bisexual — these are categories of sexual orientation.
- **Interval data** are those that are numbers where higher numbers indicate "more" of something than lower numbers.[5] In the attitude scale from the CRCC example, higher numbers indicate more support of gender equity. In a true/false knowledge test, a higher score indicates that more questions were answered correctly. In a sum of reported behaviors, a higher number indicates that more behaviors were reported as having occurred.

Before deciding what type of analysis to use, you must first decide **what data** you want to analyze and **what type of data** they are. To do this:

- Decide what you want to know.
- Identify the variables (columns) you will have to look at to find out what you want to know.
- Identify the type of data for each of the variables.

4 Those who are familiar with statistics will know that there are other types of data. However, these two types are the most common ones in surveys developed by RCCs and will be sufficient for the analyses covered in this manual.

5 Those who are familiar with statistics will know that a distinction can be drawn between data that are **interval** and data that are **ratio** based on whether the distance between consecutive numbers is always equal and whether there is an absolute zero. This distinction will not be addressed here. However, if you are working with a consultant to do more advanced analyses, they may need to make more refined distinctions between the types of data.

CRCC Data

CRCC wants to know:

- How did the students score on the pre-, post- and follow-up surveys?
- Are there changes from pre-survey scores to post-survey scores?
- Are there changes from post-survey scores to follow-up survey scores?
- Do students' attitudes differ based on gender?

Having focused on what they want to know, they identified that the variables in their spreadsheet that will answer these questions are:

- PREDATE
- POSTDATE
- FUDATE
- GENDER

They then identified the type of data these are:

- PREDATE = Interval
- POSTDATE = Interval
- FUDATE = Interval
- GENDER = Categorical

What do you want to know from your evaluation?

Which variables will answer these questions and what type of data are they?

Variable	*Type of Data*

TASK 3: IDENTIFY THE APPROPRIATE ANALYSIS

You are now ready to choose the type of analysis you will do to answer each of your evaluation questions. There are literally hundreds of statistical analyses that are used in the social sciences. This manual will teach you how to do a few very basic analyses. Some of these are as simple as **counting** the number of times an answer was given. Others are mathematical calculations you are already familiar with such as **averages**. You will also learn a few very basic ways of looking at whether or not there are meaningful **changes over time** or meaningful **differences between groups**. If there are questions you have that are not answered by these basic procedures, you may want to work with a consultant who can do more advanced analyses.

Data analysis might include counting answers, averaging, testing for change over time, or testing for differences between groups.

Before looking at the analyses that will be covered in this manual, it may be helpful to consider what we mean by **data analysis.** This phrase simply refers to different ways of **summarizing a large amount of information**. Instead of trying to make sense of the 460 answers in the CRCC worksheet plus the 60 scaled scores, data analysis will condense that information into a few numbers that can be interpreted and used.

This manual will provide step-by-step instruction on how to run five types of data analysis:

1. **Frequencies** tell us the **number of times** a particular answer occurred on the survey. For example, how many times did participants say they spoke out when they heard a sexist joke?

2. **Percentages** tell us the **proportion of times** a particular answer occurred on the survey. Percentages are especially useful when you want to compare how often an answer occurred in two or more different groups. For example, what if you want to compare how many girls spoke out when they heard a sexist joke versus how many boys spoke out? If your data includes an equal number of girls and boys you could simply compare the frequencies. However, usually we don't have the same number of people in each group. In this case, you would need to compare the percentages.

   For example, imagine you have 30 girls and 10 boys who took your survey. Among the girls, 10 girls said they spoke out against a sexist joke in the past week; 10 boys also said they did this. If you looked at the frequencies you would mistakenly think that girls and boys were the same because the same number of girls and boys reported this behavior and 10 = 10. However, if you calculated the percentage you would see a very different picture because 33% of girls said they spoke out but 100% of boys said they spoke out[6]. Therefore, a much larger proportion of boys (all of them!) had acted as engaged bystanders in this way but only one-third of the girls had.

6 Percentages are calculated by taking the frequency of the answer, dividing it by the total number of answers and multiplying that number by 100. In this example, the percentage of girls who spoke out against a sexist joke was: 10/30 = 0.33 x 100 = 33%; for boys: 10/10 = 1 x 100 = 100%.

3. **Averages** (also known as "means") describe where the data tend to gather. As long as the averages are based on the **same scale** (e.g., both are on a six-point scale or both are on a four-point scale), you can directly compare them.

   For example, imagine you had two bystander empowerment scales: one scale measured how likely people were to take action if the potential victim was a friend and the other scale measured how likely people were to act if the potential victim was a stranger. Both scales used the same set of answers that were coded from 1–5. The group's average score for a friend was 4.3, indicating that they were highly likely to intervene. But their average score for a stranger was only 2.2, indicating that they were not likely to intervene. Because these averages are on the same scale, you can directly compare these scores and conclude that this group of students is much more likely to intervene when the potential victim is a friend than a stranger.

4. **Paired samples t-tests** are used to tell us whether a group of people change their answers over time. For example, are bystander attitudes after participating in a bystander empowerment program different than they were before participating in the program. In other words, are the answers on the Post-test substantially different from the answers on the Pre-test?[7] We call this a "paired samples" test because each set of surveys (pre- and post-) are considered to be separate "samples" of data and those samples are matched up over time ("paired"). In other words, Jane's pre-survey is paired up with her post-survey, Joe's pre-survey is paired up with his post-survey, etc.

5. **Independent samples t-tests** are used to tell us whether two separate ("independent") groups give different answers. For example, are bystander attitudes of girls different than bystander attitudes of boys? While we could compare their percentages, the t-test gives us additional information by telling us whether the differences are meaningful. Specifically, is the difference substantially larger than what we would expect to occur by chance. This is an important distinction. For example, what if the girls have an average score of 4.2 and the boys have an average score of 4.3. Do we really want to say that boys are more likely to intervene than girls are when the difference is so small? The t-test will give us a consistent way of making that determination.

7 Those who have a background in statistics or who have simply read evaluation reports may be familiar with the term "statistical significance" or "a statistically significant difference." The full, technical definition of statistical significance is beyond the scope of this manual so we will rely on a simplified, conceptual definition of whether the difference is substantial or meaningful. This will be discussed more in the section on Interpreting Quantitative Data.

The type of analysis you use depends on what you want to know and the type of data you have. In other words, the type of analysis is the flashlight. In order to use it, you have to have the right size battery (i.e., the right type of data). The table below summarizes how these different factors fit together.

IF YOU WANT TO KNOW...	AND YOUR DATA ARE...	RUN THIS TEST...
How many people gave a certain answer	Categorical OR Interval	Frequency
What portion of people gave a certain answer (in total or by categories)	Categorical OR Interval	Percentage
What the center point is around which the answers tend to fall	Interval	Average
If there are differences in the **same** group over time	Interval	Paired samples t-test[8]
If there are differences **between** groups	Interval[9]	Independent samples t-test

Warning: **Paired and independent samples t-tests require data from at least 30 people**. If you have data from less than 30 people different tests must be used.[10]

You now have a worksheet with clean data. You have identified the types of variables you have and you have identified the type of analysis that will answer your evaluation questions. So now you are ready to run the data analysis(es).

8 Those who are familiar with statistics will know that both paired and independent samples t-tests assume that the data are normally distributed. For the purposes of this resource kit, we will assume this assumption is met. If you are working with a consultant, they should verify that the assumption is met and, if it is not, may wish to use a different type of analysis.

9 If your data are categorical you will probably want to run a chi-squared analysis. This technique is not covered in this manual but can be done with Microsoft Excel. There are also other tests possible depending on the exact nature of the categorical data.

10 To make the illustrations in this manual easier to read, data for only 20 people were entered. This was done solely for illustrative purposes. Pretend that there were at least 10 more surveys in the worksheet.

TASK 4: RUN THE ANALYSIS

Before you start running analyses you might want to copy and paste your data into a new worksheet. If you accidentally delete or change something that you didn't want to, you can go back to your original survey worksheet and will not have to re-enter any data. From this point forward you will see in the worksheet illustrations that in the CRCC example there are separate worksheets for the analyses.

Note: The instructions that follow will suggest that you put functions, graphs and results in certain locations on your worksheet. However, you can put them wherever you want. The suggestions made in this manual are based on what many people find to be easy. **What matters is that you label and save data in the way that makes sense to you**. If you can find your data and results and you are clear as to what number is what, then any way you organize your data is fine.

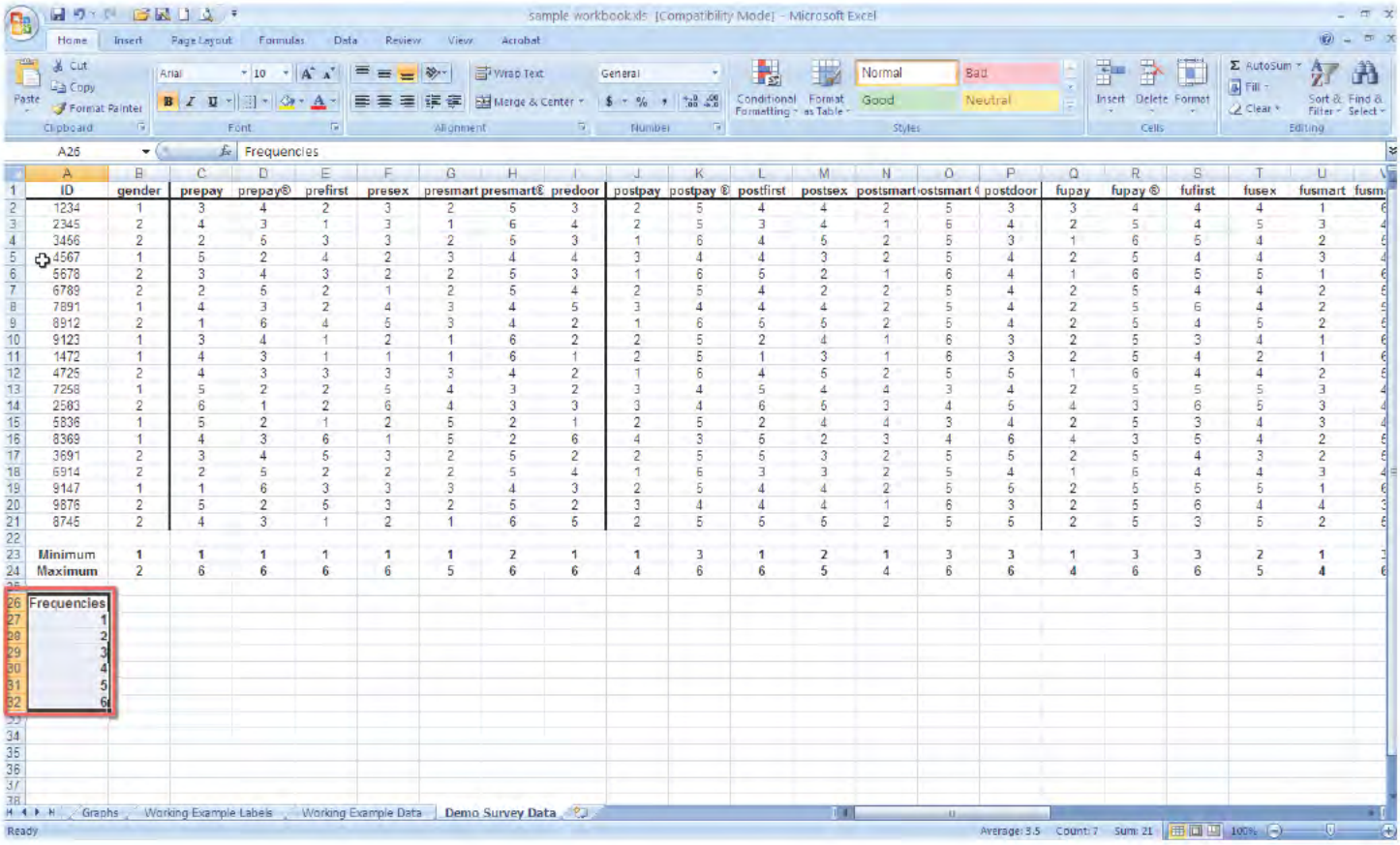
Frequencies

Frequencies tell us how often participants gave a particular answer.

There are multiple ways of calculating frequencies in Excel. The function demonstrated here is the "count if" function. In this approach you will tell the computer to count up how many times a particular answer was given.

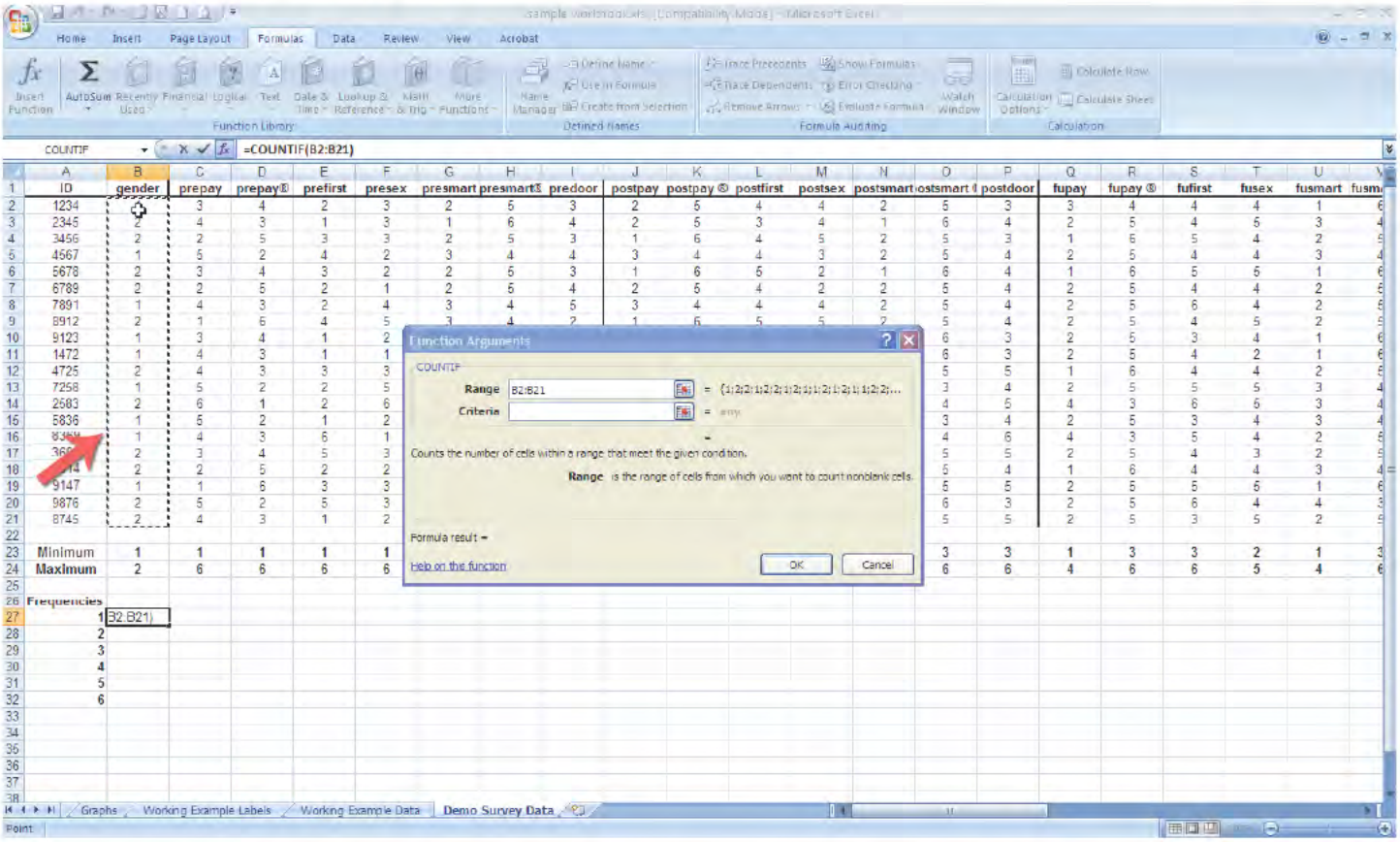
To calculate frequencies:

- Before you run the function, label where your frequencies will go. It is important to label your results from the beginning. The results will appear on the worksheet simply as numbers. Without labels, you may quickly forget to what the numbers refer.
- As shown in the picture on the next page, it is recommended that you put frequencies (and percentages) at the bottom of each question's column and underneath where you already calculated the minimum and maximum values for each question. Give each possible answer its own row.



- To fill in the **frequencies,** you will use Excel's functions much like you did when you calculated the minimum and maximum scores. However, you will be using the **COUNTIF function**.
- Click to put your cursor in the cell where you want the first frequency result to go. In the CRCC example, it would be in column B (gender) row 27 (frequency: 1).
- In the **Formula ribbon**, click on the **Insert Function button**. This will open up the Function Dialog Box.
- In the Function Dialog Box make sure that where it says "Or select a category" the category is **"Statistical."** If it is not, then use the drop-down list to select the Statistical functions.
- In the "Select a function" box scroll up until you find the **COUNTIF** function. Highlight COUNTIF and **click OK.**

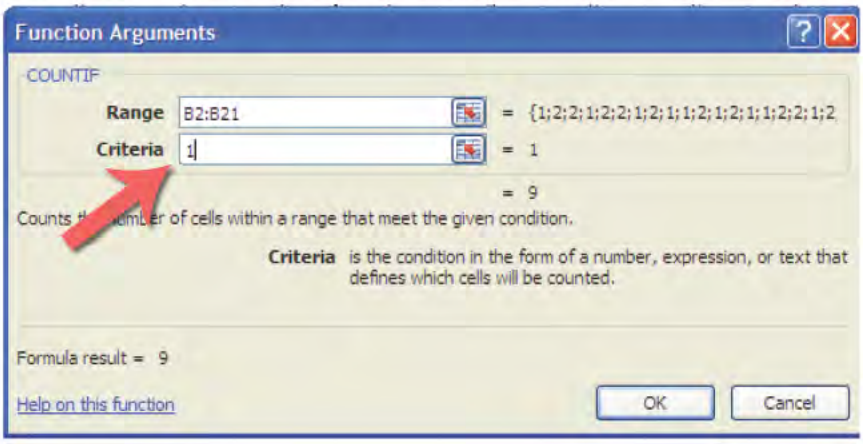
- After you click OK, a dialog box will open up. The **range box** tells you on what cells the function will be performed. The cells are referred to by their column letter and row number. The colon in-between the cells means that Excel will look at all of the cells in between those two cells. You can change the cell range by typing in the cell range OR by using your mouse to highlight the cells you want included.



Example: dashed line shows that the range includes cells B2 through B21

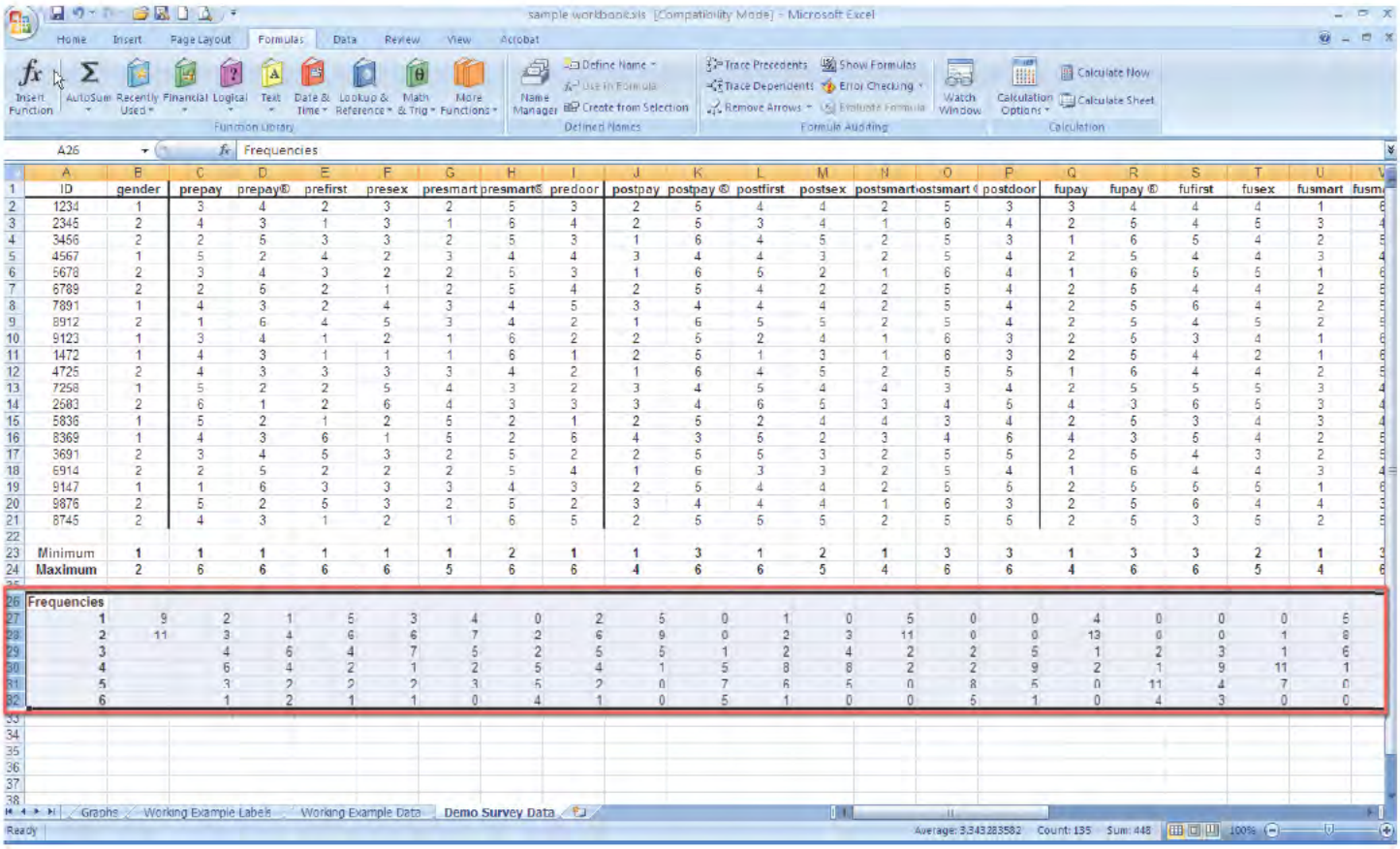
Example: Range = B2:B21

- In the **Criteria box** you need to type in which **answer you want counted.**
- **Click OK.** You will see that the number of times "1" was given for "gender" now appears.



- Once you have calculated the first frequency for the first question, you can automatically repeat that function for all of the remaining questions. You can do this by holding your mouse so the cursor hovers over the lower right corner of that cell where you have the first frequency. You will see that a + sign appears. Hold down on the mouse and drag that + sign across the row, highlighting all of the cells as you go. When you let go, the frequencies for each question will appear in the appropriate places.
- **Repeat** these steps to calculate the frequencies for the remaining answers.

In the CRCC example, once all frequency scores were completed they appeared as illustrated below:



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	ID	gender	prepay	prepay®	prefirst	presex	presmart	presmart®	predoor	postpay	postpay ©	postfirst	postsex	postsmart	postsmart	postdoor	fupay	fupay ®	fufirst	fusex	fusmart
2	1234	1	3	4	2	3	2	5	3	2	5	4	4	2	5	3	3	4	4	4	1
3	2345	2	4	3	1	3	1	6	4	2	5	3	4	1	6	4	2	5	4	5	3
4	3456	2	2	5	3	3	2	5	3	1	6	4	5	2	5	3	1	6	5	4	2
5	4567	1	5	2	4	2	3	4	4	3	4	4	3	2	5	4	2	5	4	4	3
6	5678	2	3	4	3	2	2	5	3	1	6	5	2	1	6	4	1	6	5	5	1
7	6789	2	2	5	2	1	2	5	4	2	5	4	2	2	5	4	2	5	4	4	2
8	7891	1	4	3	2	4	3	4	5	3	4	4	4	2	5	4	2	5	6	4	2
9	8912	2	1	6	4	5	3	4	2	1	6	5	5	2	5	4	2	5	4	5	2
10	9123	1	3	4	1	2	1	6	2	2	5	2	4	1	6	3	2	5	3	4	1
11	1472	1	4	3	1	1	1	6	1	2	5	1	3	1	6	3	2	5	4	2	1
12	4725	2	4	3	3	3	3	4	2	1	6	4	5	2	5	5	1	6	4	4	2
13	7258	1	5	2	2	5	4	3	2	3	4	5	4	4	3	4	2	5	5	5	3
14	2583	2	6	1	2	6	4	3	3	3	4	6	5	3	4	5	4	3	6	5	3
15	5836	1	5	2	1	2	5	2	1	2	5	2	4	4	3	4	2	5	3	4	3
16	8369	1	4	3	6	1	5	2	6	4	3	5	2	3	4	6	4	3	5	4	2
17	3691	2	3	4	5	3	2	5	2	2	5	5	3	2	5	5	2	5	4	3	2
18	6914	2	2	5	2	2	2	5	4	1	6	3	3	2	5	4	1	6	4	4	3
19	9147	1	1	6	3	3	3	4	3	2	5	4	4	2	5	5	2	5	5	5	1
20	9876	2	5	2	5	3	2	5	2	3	4	4	4	1	6	3	2	5	6	4	4
21	8745	2	4	3	1	2	1	6	5	2	5	5	5	2	5	5	2	5	3	5	2
22																					
23	Minimum	1	1	1	1	1	1	2	1	1	3	1	2	1	3	3	1	3	3	2	1
24	Maximum	2	6	6	6	6	5	6	6	4	6	6	5	4	6	6	4	6	6	5	4
26	Frequencies																				
27	1	9	2	1	5	3	4	0	2	5	0	1	0	5	0	0	4	0	0	0	5
28	2	11	3	4	6	6	7	2	6	9	0	2	3	11	0	0	13	0	0	1	8
29	3		4	6	4	7	5	2	5	5	1	2	4	2	2	5	1	2	3	1	6
30	4		6	4	2	1	2	5	4	1	5	8	8	2	2	9	2	1	9	11	1
31	5		3	2	2	2	3	5	2	0	7	6	5	0	8	5	0	11	4	7	0
32	6		1	2	1	1	0	4	1	0	5	1	0	0	5	1	0	4	3	0	0

Percentages

A percentage tells us what proportion of the participants gave a particular answer. It lets us compare groups that aren't the same size.

A **percentage** is calculated by dividing the number of people who gave a particular answer (for example, the number of people who said they were female) by the total number of people you have data for (for example, the total number of people who took the survey). This will give you a decimal point percentage. To convert to a whole number percentage, simply multiple your answer by 100.

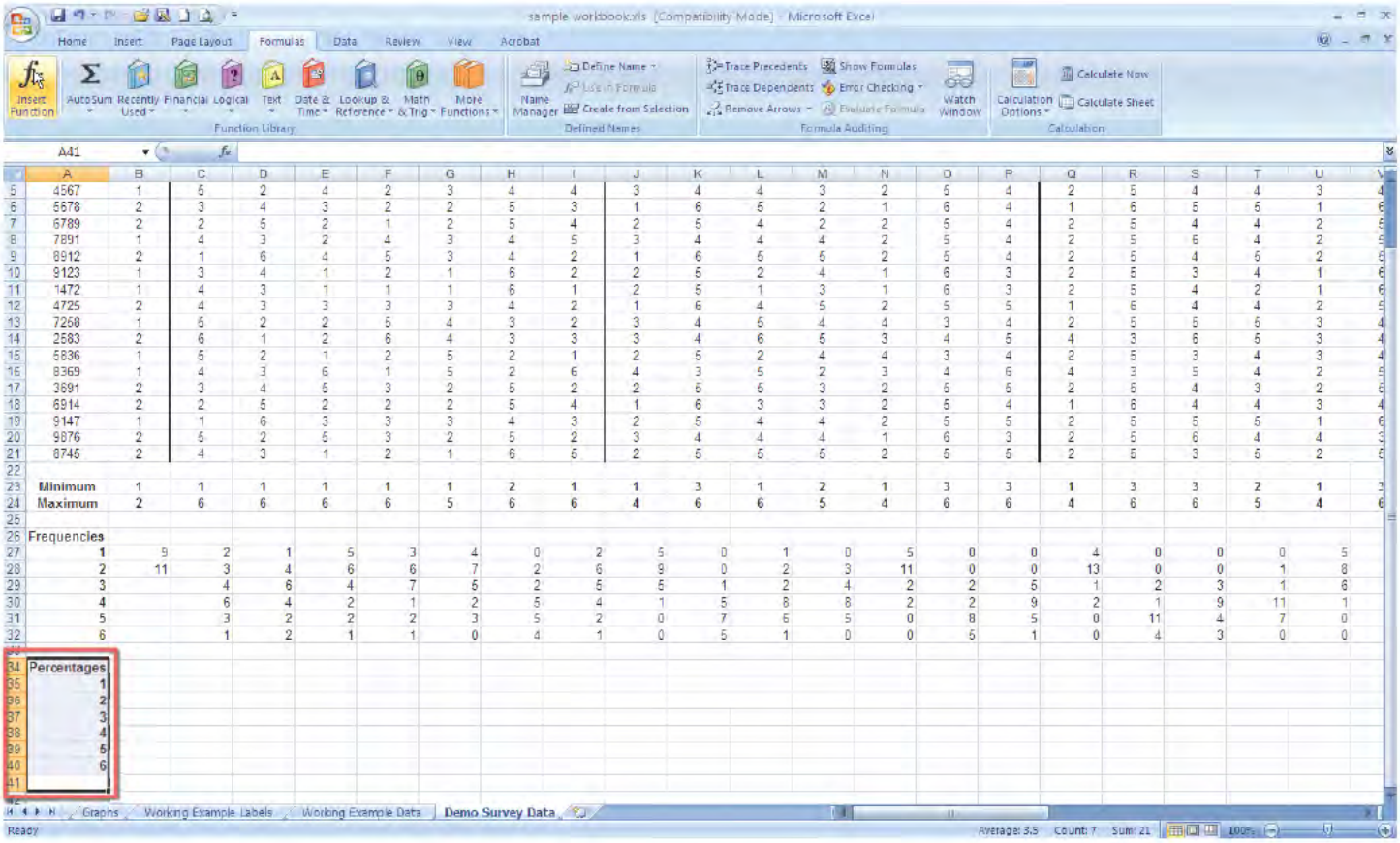
In the CRCC example we see that there were nine people who said they were female out of 20 total people who took the survey. Therefore, the percentage of females is:

9 divided by 20 = 0.45

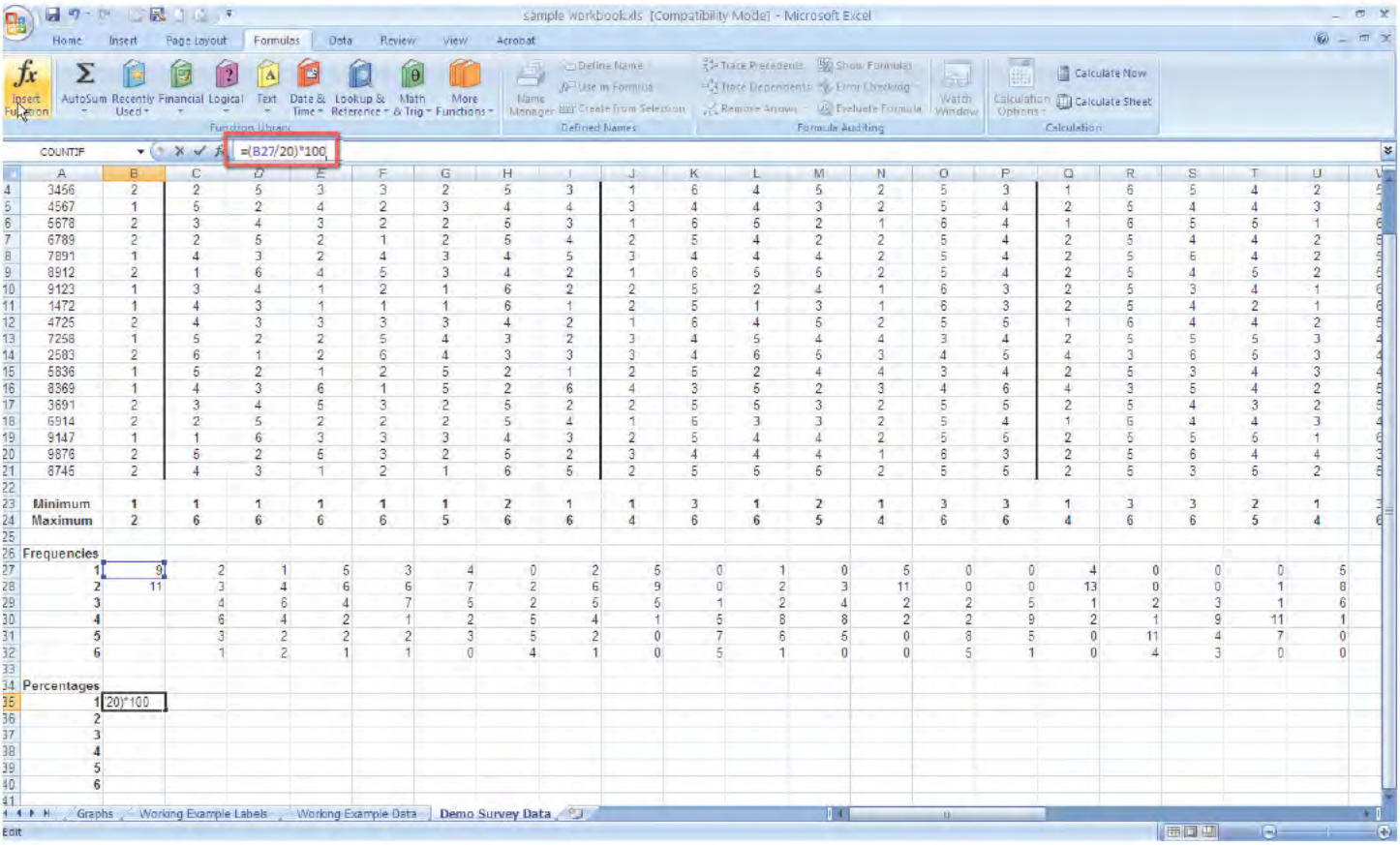
0.45 x 100 = 45%

The quickest way to calculate percentages in Microsoft Excel is to write the frequencies you have already calculated into a percentage formula.

- *To do this:* Before you run the function, **label where your percentages will go.** It is important to label your results from the beginning. The results will appear on the worksheet simply as numbers. Without labels you may quickly forget to what the numbers refer.
- It is recommended that you put percentages at the bottom of each question's column and **underneath where you already calculated the frequencies**. Just like with the frequencies, give each possible answer its own row.

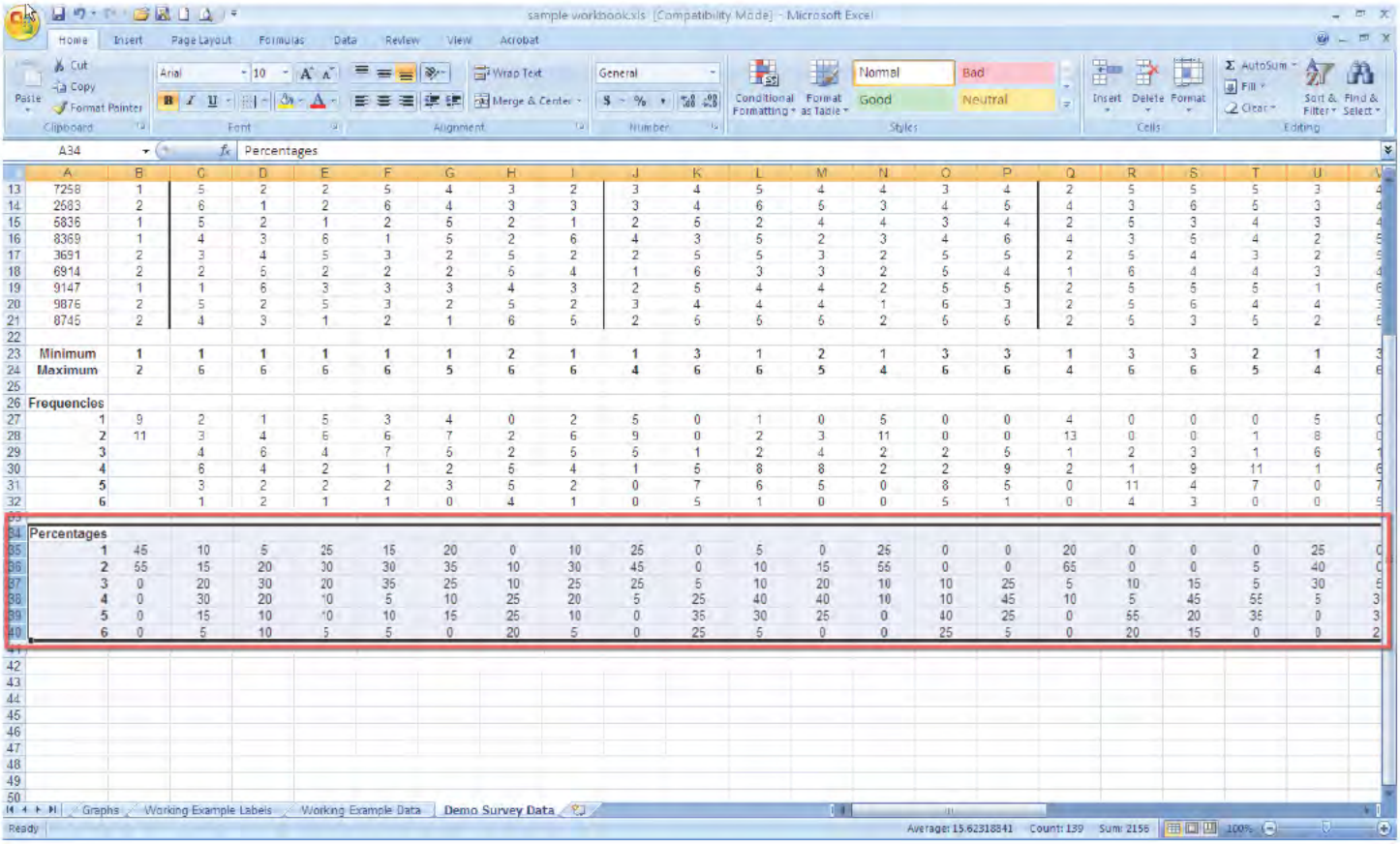


- To fill in the **percentages,** you will **write the formula in the function box** that appears immediately above your data.
  - **Click on the cell** where you want the first percentage result to go. In the CRCC example it would be in column B (gender) row 35 (frequency: 1).
  - Type in the formula for calculating a percentage. For the CRCC example the first formula looks like this: **=(B27/20)*100.**
    - Every formula begins with an **equal sign.**
    - The **cell** that has the **frequency** you need comes next.
    - The forward slash / symbol means "divided by."
    - The next number is the **total** people for which you have data.
    - The parentheses **( )** must go around those two numbers. This tells the computer to divide those two numbers before doing anything else.
    - The asterisk * means "multiply by."
    - The last number is **100** to change your decimal answer into a whole number.
    - So in this case, we are telling the computer: "Take the frequency from cell B27 and divide it by 20. Then multiply that answer by 100."



- Once you have calculated the first percentage for the first question, you can automatically repeat that function for all of the remaining questions. You can do this by holding your mouse so the cursor hovers over the lower right corner of that cell where you have the first percentage. You will see that a + sign appears. Hold down on the mouse and drag that + sign across the row, highlighting all of the cells as you go. When you let go, the percentages for each question will appear in the appropriate places.
- **Repeat** these steps to calculate the percentages for the remaining answers.

In the CRCC example, once all percentage scores were calculated they appeared as illustrated below:



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
13	7258	1	5	2	2	5	4	3	2	3	4	5	4	4	3	4	2	5	5	5	3
14	2683	2	6	1	2	6	4	3	3	3	4	6	5	3	4	5	4	3	6	5	3
15	5836	1	5	2	1	2	5	2	1	2	5	2	4	4	3	4	2	5	3	4	3
16	8369	1	4	3	6	1	5	2	6	4	3	5	2	3	4	6	4	3	5	4	2
17	3691	2	3	4	5	3	2	5	2	2	5	5	3	2	5	5	2	5	4	3	2
18	6914	2	2	5	2	2	2	6	4	1	6	3	3	2	5	4	1	6	4	4	3
19	9147	1	1	6	3	3	3	4	3	2	5	4	4	2	5	5	2	5	5	5	1
20	9876	2	5	2	5	3	2	5	2	3	4	4	4	1	6	3	2	5	6	4	4
21	8745	2	4	3	1	2	1	6	5	2	5	5	5	2	5	5	2	5	3	5	2
22																					
23	**Minimum**	**1**	**1**	**1**	**1**	**1**	**1**	**2**	**1**	**1**	**3**	**1**	**2**	**1**	**3**	**3**	**1**	**3**	**3**	**2**	**1**
24	**Maximum**	**2**	**6**	**6**	**6**	**6**	**5**	**6**	**6**	**4**	**6**	**6**	**5**	**4**	**6**	**6**	**4**	**6**	**6**	**5**	**4**
25																					
26	**Frequencies**																				
27	1	9	2	1	5	3	4	0	2	5	0	1	0	5	0	0	4	0	0	0	5
28	2	11	3	4	6	6	7	2	6	9	0	2	3	11	0	0	13	0	0	1	8
29	3		4	6	4	7	5	2	5	5	1	2	4	2	2	5	1	2	3	1	6
30	4		6	4	2	1	2	5	4	1	5	8	8	2	2	9	2	1	9	11	1
31	5		3	2	2	2	3	5	2	0	7	6	5	0	8	5	0	11	4	7	0
32	6		1	2	1	1	0	4	1	0	5	1	0	0	5	1	0	4	3	0	0
33																					
34	**Percentages**																				
35	1	45	10	5	25	15	20	0	10	25	0	5	0	25	0	0	20	0	0	0	25
36	2	55	15	20	30	30	35	10	30	45	0	10	15	55	0	0	65	0	0	5	40
37	3	0	20	30	20	35	25	10	25	25	5	10	20	10	10	25	5	10	15	5	30
38	4	0	30	20	10	5	10	25	20	5	25	40	40	10	10	45	10	5	45	55	5
39	5	0	15	10	10	10	15	25	10	0	35	30	25	0	40	25	0	55	20	35	0
40	6	0	5	10	5	5	0	20	5	0	25	5	0	0	25	5	0	20	15	0	0
41																					
42																					
43																					
44																					
45																					
46																					
47																					
48																					
49																					
50																					

We said earlier that the point of data analysis is to **summarize the data**. We want to reduce the answers to a few statistics that can capture the bigger picture. Up until this point, the data analysis we have done has still focused on individual questions. In the next section of this manual, we will talk about how to interpret those detailed results in a useful way.

For the remaining analyses (averages, paired samples and independent samples t-tests) we will intentionally shift our focus to the summary level. For the remaining analyses we will work **only with the scaled scores** that were calculated earlier. While you can perform any of the following analyses on individual questions, the example shown here will focus on the scaled scores as a way of looking at the bigger picture.

Averages

Averages summarize what, overall, the scores were like.

We will use averages to answer the question, "On average, what were their attitudes toward dating and heterosexual relationships on each of the surveys?"

Before you start running analyses you might want to copy and paste your **scaled scores** into a new worksheet. This will do two things:

1. It will remove all of the question-level data and analyses from your view. There is no need to be distracted by question-level data when you are working with the scaled scores.
2. If you accidentally delete or change something that you didn't want to, you can go back to your original survey worksheet and will not have to re-enter any data. From this point forward you will see in the worksheet illustrations for the CRCC example that there is a separate worksheet for the averages and t-tests.

To copy your scaled scores:

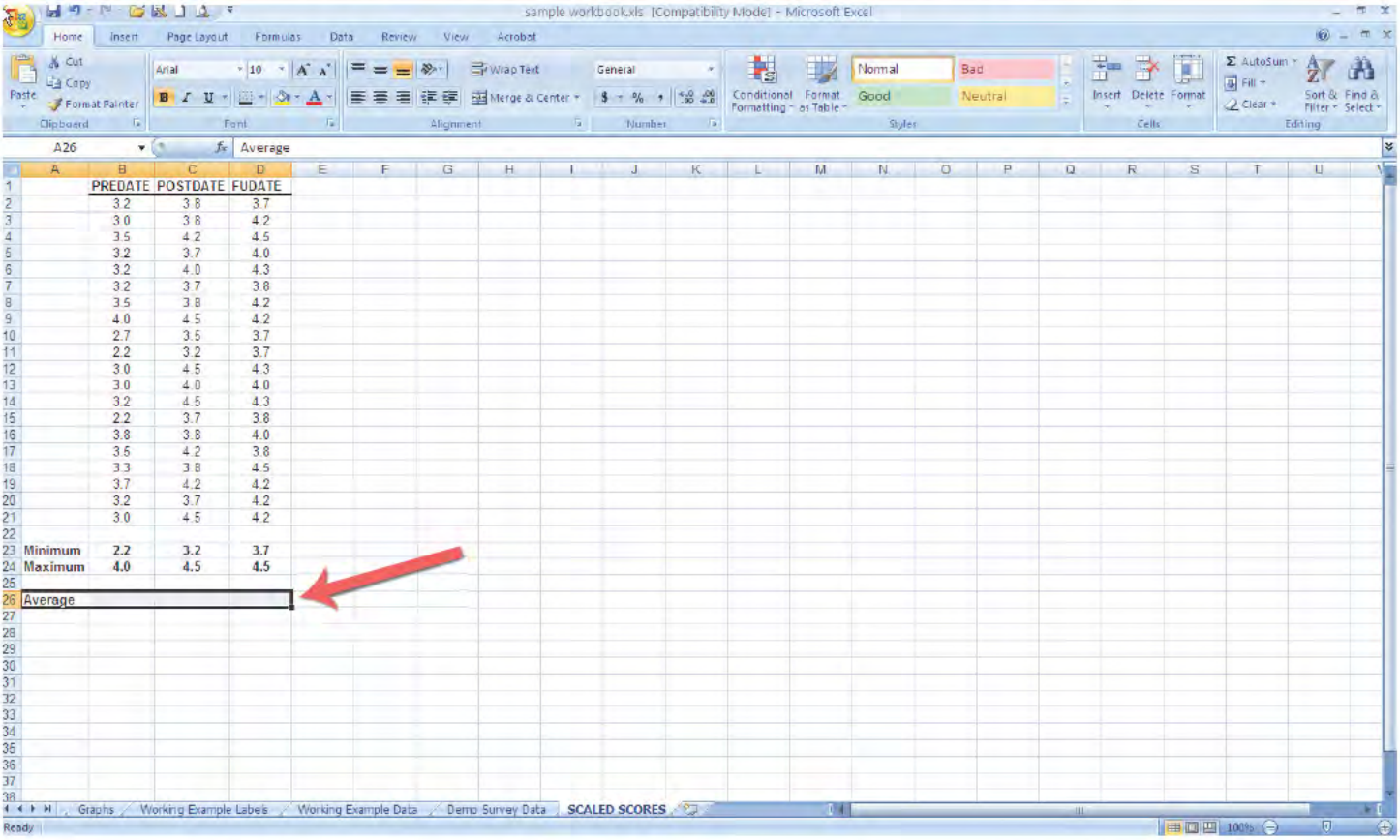
- Use your mouse to highlight the scaled scores from one of your previous worksheets.
- Use the Edit menu to select Copy OR press CTRL+C on your keyboard.
- Click on the next worksheet tab at the bottom of the screen (or insert a new worksheet).
- Click on the first row of the second column in the blank worksheet. You will probably want to use the first column to label your analyses.
- Because the cells you copied included formulas, you will now need to do a **Special Paste**. Use the Edit menu to select **Special Paste**.
- That will open up a dialog box that will ask you what you want to paste. In the Paste list select **Values**.
  - Click OK.
  - You will see that the values (i.e., the scaled scores) have now been pasted on your new worksheet.

    Note: If you see "REF#" instead of the scaled score values, this means that you did not do a special paste with only the values. Instead, you tried to paste the formulas. Because the formulas refer to the cells on your prior worksheets, they will not work on your new worksheet. Try the paste again, making sure that you select **Special Paste** from the Edit menu.

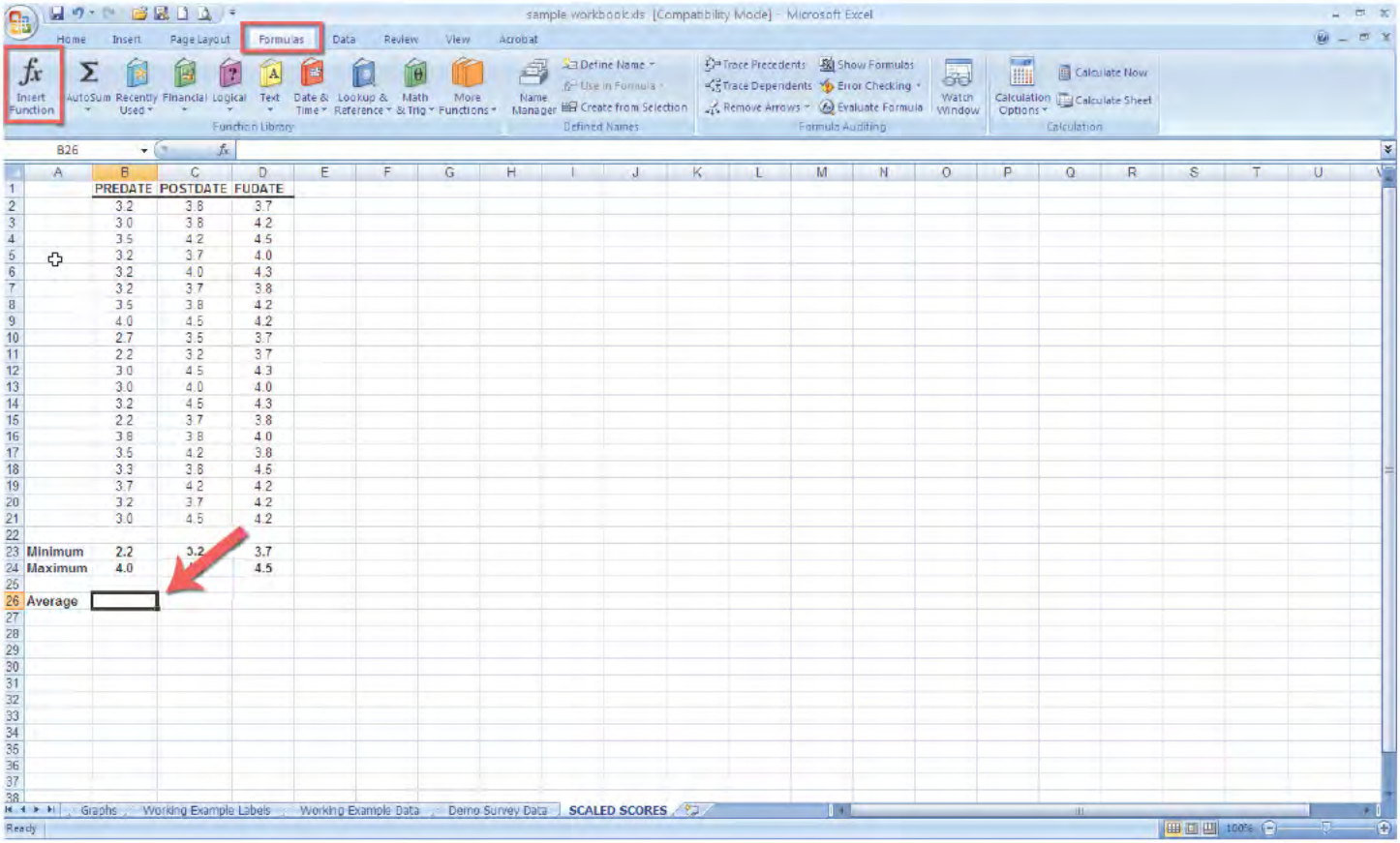
We will use Microsoft Excel's functions to automatically calculate the averages. We will use the AVERAGE function.

To calculate averages:

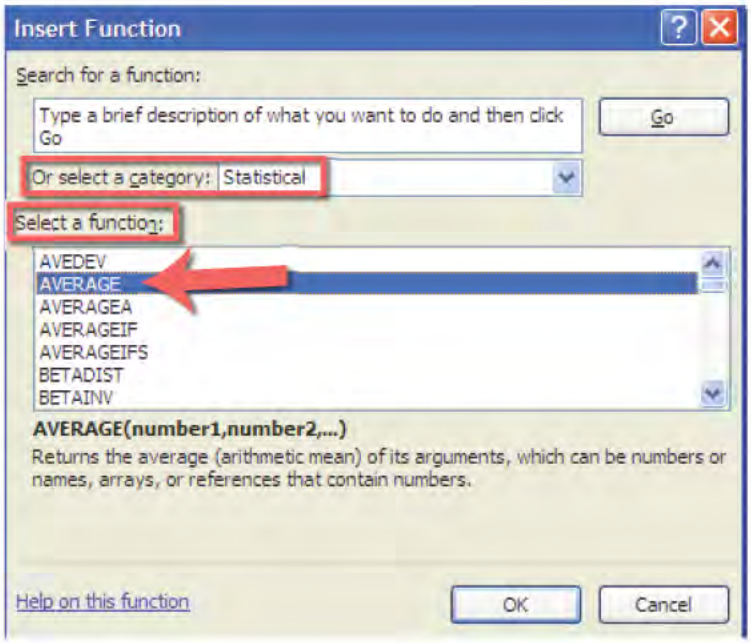
- Before you run the function, **label where your averages will go.** It is important to label your results from the beginning. The results will appear on the worksheet simply as numbers. Without labels you may quickly forget to what the numbers refer. As shown in the picture below, it is recommended that you put averages at the **bottom** of each subscale's column.



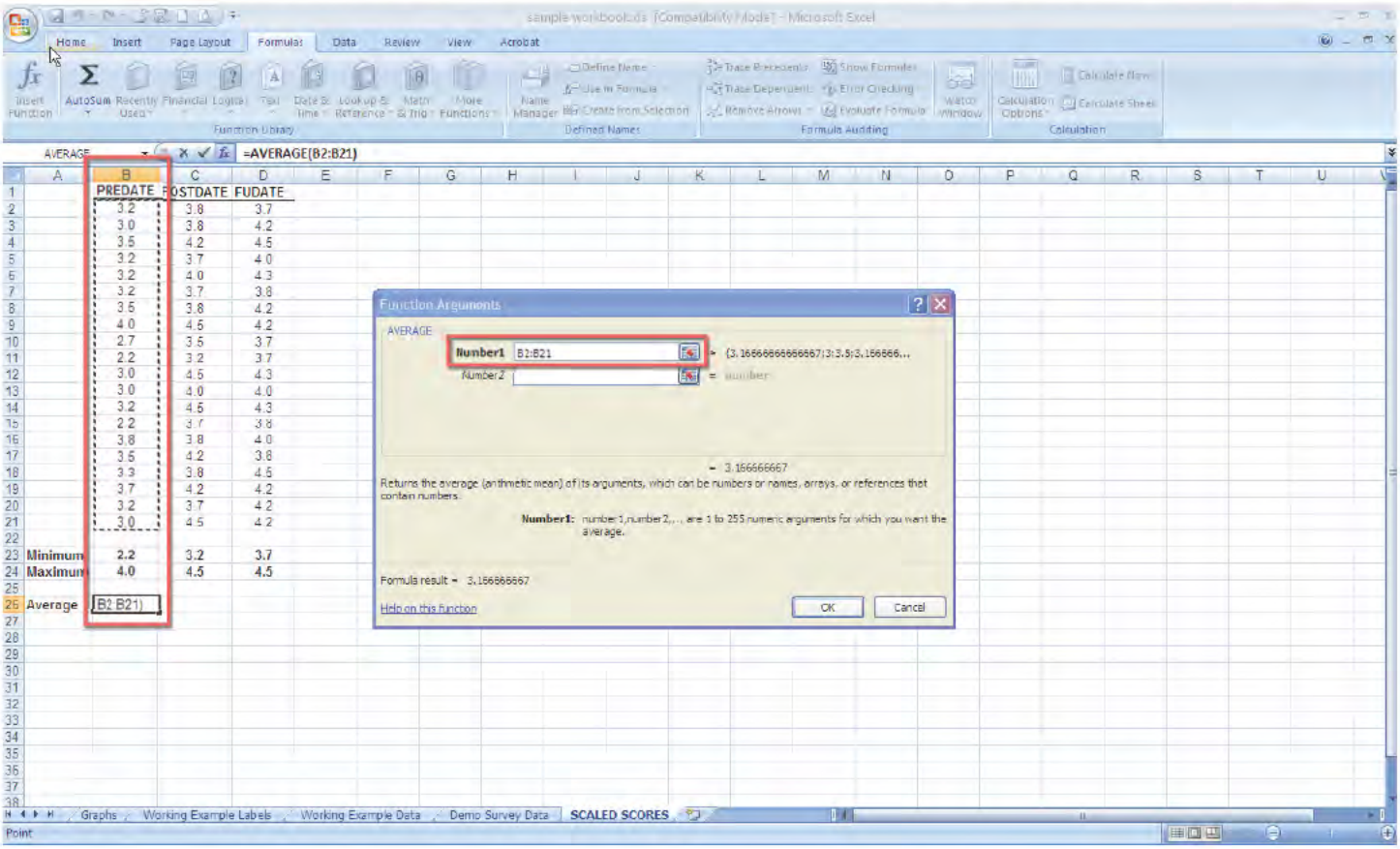
- To fill in the **averages,** you will use Excel's functions much like you did when you calculated the minimum and maximum scores. However, you will be using the **AVERAGE** function.
  - **Click** to put your cursor in the cell where you want the first average result to go. (In the CRCC example it would be in column B (predate) row 26.)
  - In the **Formula ribbon**, click on the **Insert Function button**. This will open up the Function Dialog Box.



- In the Function Dialog Box, make sure that where it says "Or select a category" the category is "**Statistical.**" If it is not, then use the drop-down list to select the Statistical functions.
- In the "**Select a function**" box, scroll until you find the **AVERAGE** function. Highlight AVERAGE and click OK. *Warning*: There are multiple functions for averaging. Make sure you are using the plain AVERAGE function. Do not use AVEDEV or AVERAGE A. These are different functions and will give you different results.
- Click OK.



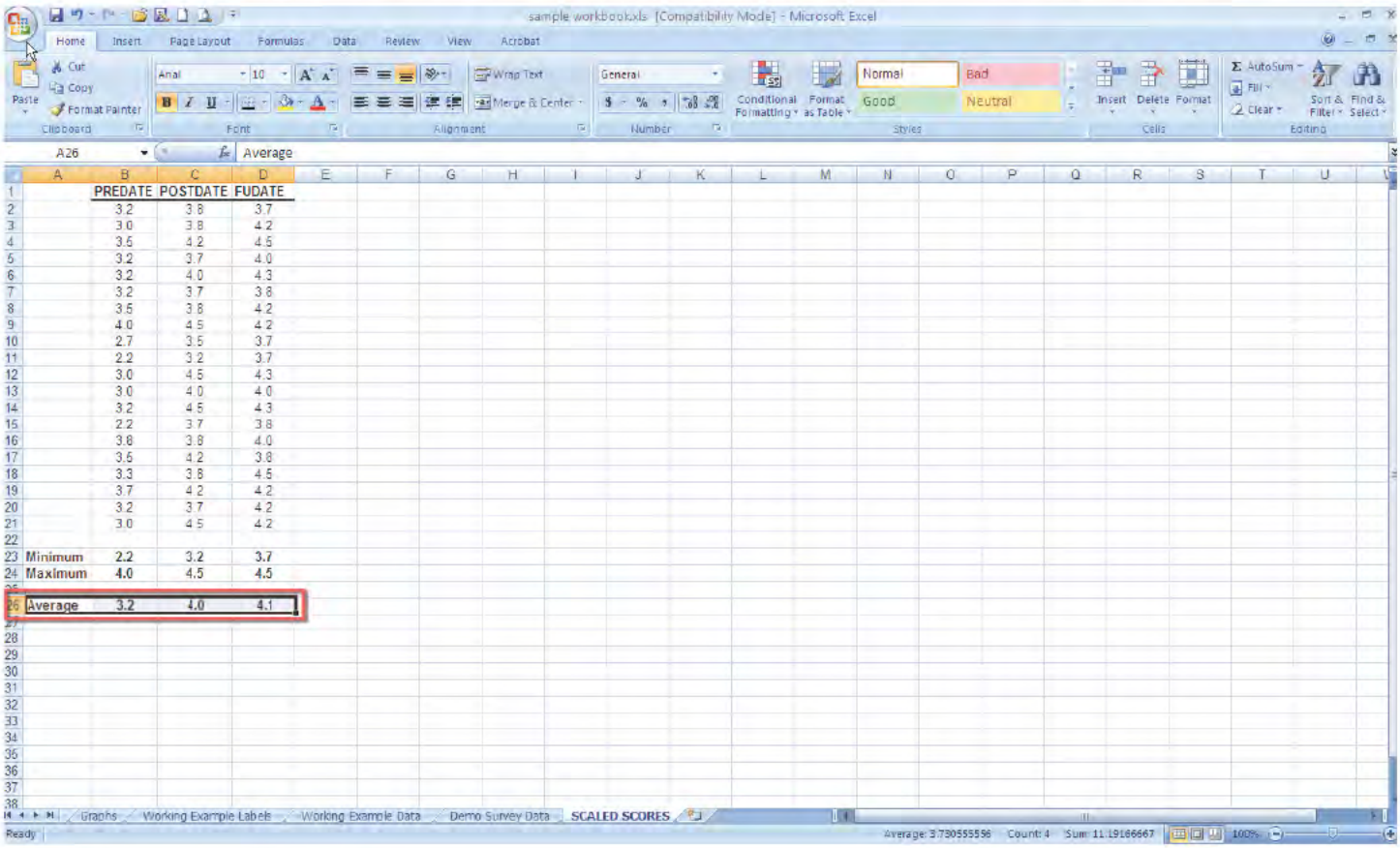
- After you click OK a dialog box will open up. The **range box** tells you on what cells the function will be performed. The cells are referred to by their column letter and row number. The colon in-between the cells means that Excel will look at all of the cells in between those two cells. You can change the cell range by typing in the cell range OR by using your mouse to highlight the cells you want included as shown in the picture below.



- **Click OK.** You will see that the average for column now appears.

- Once you have calculated the first average, you can automatically repeat that function for all of the remaining scales. You can do this by holding your mouse so the cursor hovers over the lower right corner of that cell where you have the first average. You will see that a + sign appears. Hold down on the mouse and drag that + sign across the row, highlighting all of the cells as you go. When you let go, the averages for each scale will appear in the appropriate places.

In the CRCC example, once all average scores were completed they appeared as illustrated below:



	PREDATE	POSTDATE	FUDATE
	3.2	3.8	3.7
	3.0	3.8	4.2
	3.5	4.2	4.5
	3.2	3.7	4.0
	3.2	4.0	4.3
	3.2	3.7	3.8
	3.5	3.8	4.2
	4.0	4.5	4.2
	2.7	3.5	3.7
	2.2	3.2	3.7
	3.0	4.5	4.3
	3.0	4.0	4.0
	3.2	4.5	4.3
	2.2	3.7	3.8
	3.8	3.8	4.0
	3.5	4.2	3.8
	3.3	3.8	4.5
	3.7	4.2	4.2
	3.2	3.7	4.2
	3.0	4.5	4.2
Minimum	2.2	3.2	3.7
Maximum	4.0	4.5	4.5
Average	3.2	4.0	4.1

Paired Samples T-test

Now that we know what the average scores are for each of the surveys, we can look at them and see how they changed over time. However, sometimes these changes are very small.

For example, in the CRCC illustration on the previous page we see that the average scores were:

Pre-survey = 3.2

Post-survey = 4.0

Follow-up survey = 4.1

We see that the averages did go up each time. But is the amount they went up a little or a lot? We need a consistent and objective way of deciding if these differences represent a meaningful (or "real") change. To do this we will use the **paired samples t-test**. This test looks at **how much each individual person's changed** from one survey to the next. It determines if those changes were substantial (or "significant") or if they were simply the kind of random fluctuations we would expect even if the prevention program was not working.

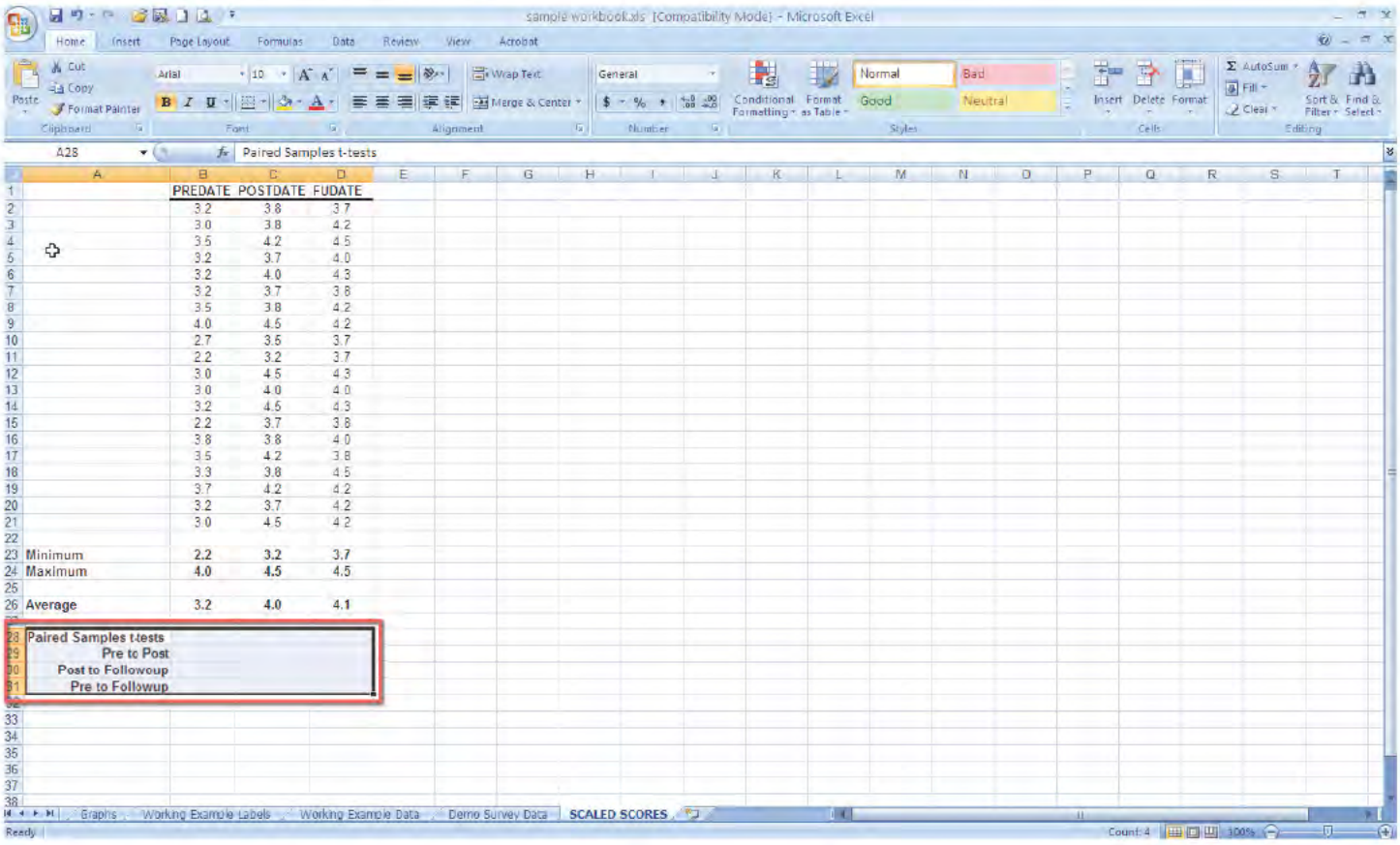
To run a paired samples t-test:

- Before you run the function, **label where your test results will go.** It is important to label your results from the beginning. The results will appear on the worksheet simply as numbers. Without labels you may quickly forget what the numbers refer to.
- As shown in the picture below, it is recommended that you put t-test results **below** the averages.
- The number of paired samples t-tests you run will depend on how many comparisons you are making. Each comparison is one test. For example, CRCC has three comparisons to make:

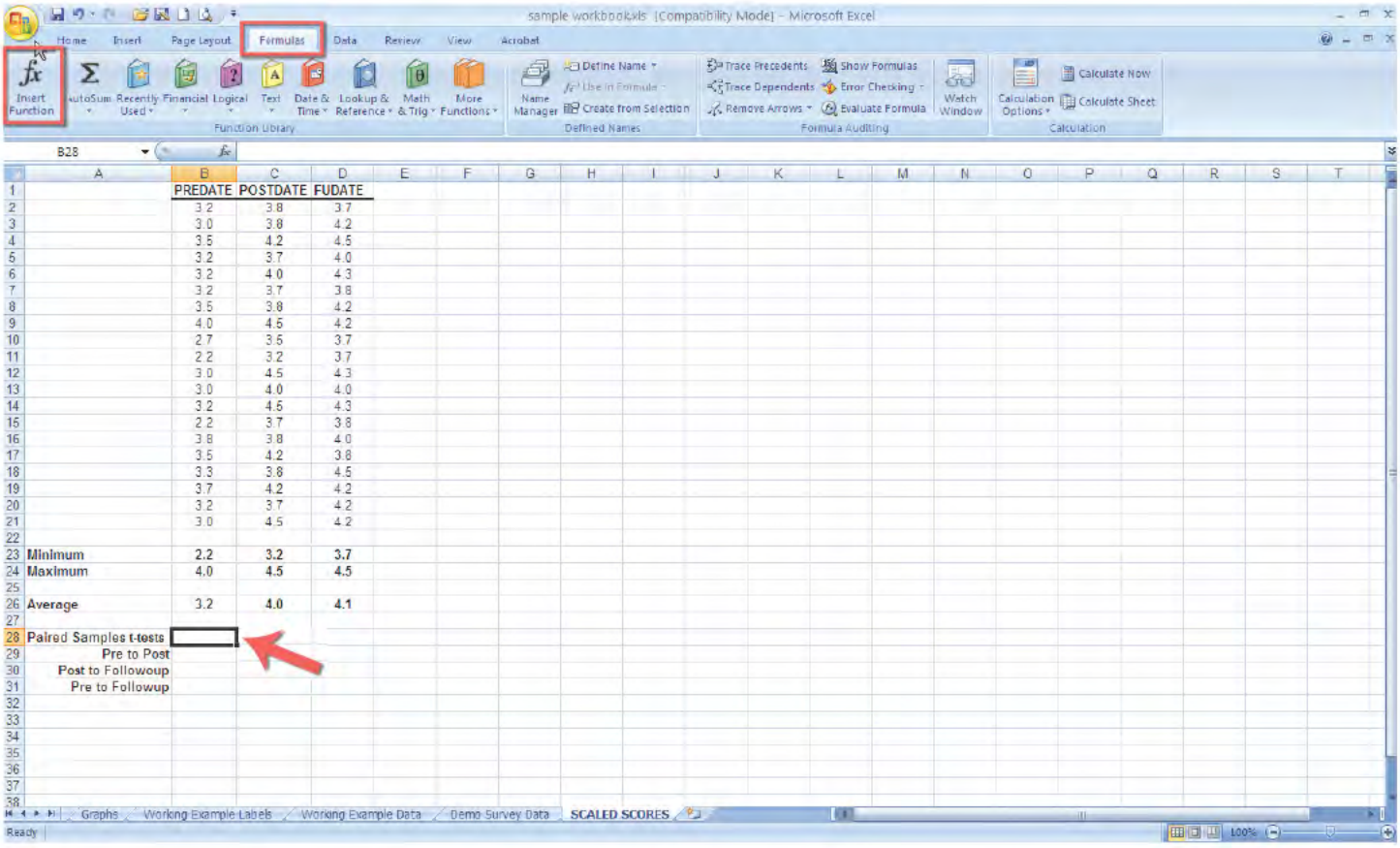
  1. **Pre-survey vs. Post-survey** tells them if there was a change from before the prevention program to after.

  2. **Post-survey vs. Follow-up survey** tells them if the effects of the program were sustained over time.

  3. **Pre-survey vs. Follow-up survey** tells them if, over time, the effects of the program went back to pre-program levels or were still higher than before the program (even if they are lower than immediately after the program ended).

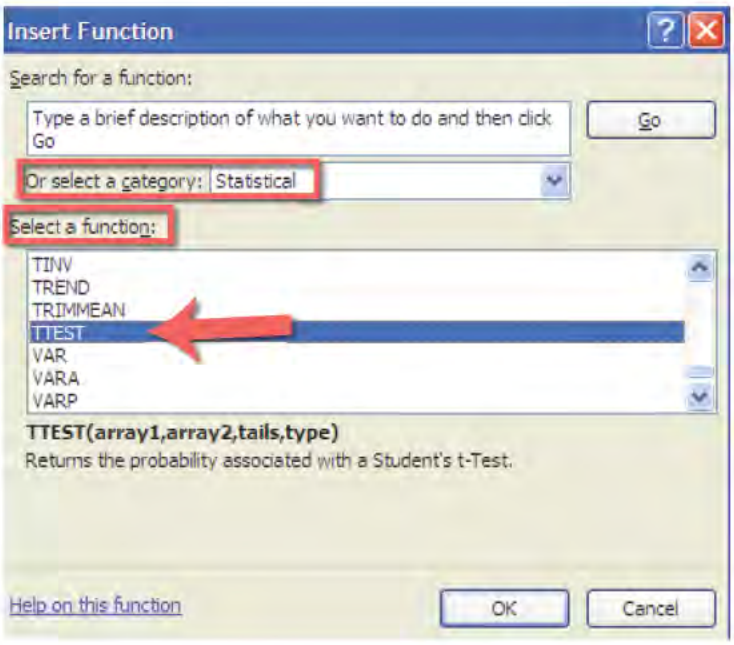


- To calculate the **paired t-test** results you will use Excel's functions like you did when you calculated the minimum and maximum scores and when you calculated averages. However, you will be using the TTEST function.
- **Click** to put your cursor in the cell where you want the first t-test result to go. In the CRCC example, it would be in column B, row 28.
  - In the **Formula ribbon**, click on the **Insert Function button**. This will open up the Function Dialog Box.



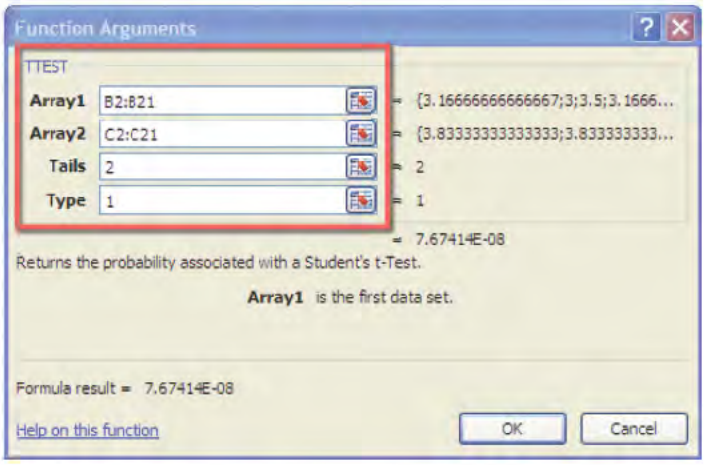
	A	B	C	D
1		PREDATE	POSTDATE	FUDATE
2		3.2	3.8	3.7
3		3.0	3.8	4.2
4		3.5	4.2	4.5
5		3.2	3.7	4.0
6		3.2	4.0	4.3
7		3.2	3.7	3.8
8		3.5	3.8	4.2
9		4.0	4.5	4.2
10		2.7	3.5	3.7
11		2.2	3.2	3.7
12		3.0	4.5	4.3
13		3.0	4.0	4.0
14		3.2	4.5	4.3
15		2.2	3.7	3.8
16		3.8	3.8	4.0
17		3.5	4.2	3.8
18		3.3	3.8	4.5
19		3.7	4.2	4.2
20		3.2	3.7	4.2
21		3.0	4.5	4.2
22				
23	Minimum	2.2	3.2	3.7
24	Maximum	4.0	4.5	4.5
25				
26	Average	3.2	4.0	4.1
27				
28	Paired Samples t-tests			
29	Pre to Post			
30	Post to Followoup			
31	Pre to Followup			
32				
33				
34				
35				
36				
37				
38				

- In the Function Dialog Box make sure that where it says "Or select a category" the category is "**Statistical.**" If it is not, use the drop-down list to select the Statistical functions.
- In the "Select a function" box scroll until you find the **TTEST** function. Highlight TTEST and then **click OK.**

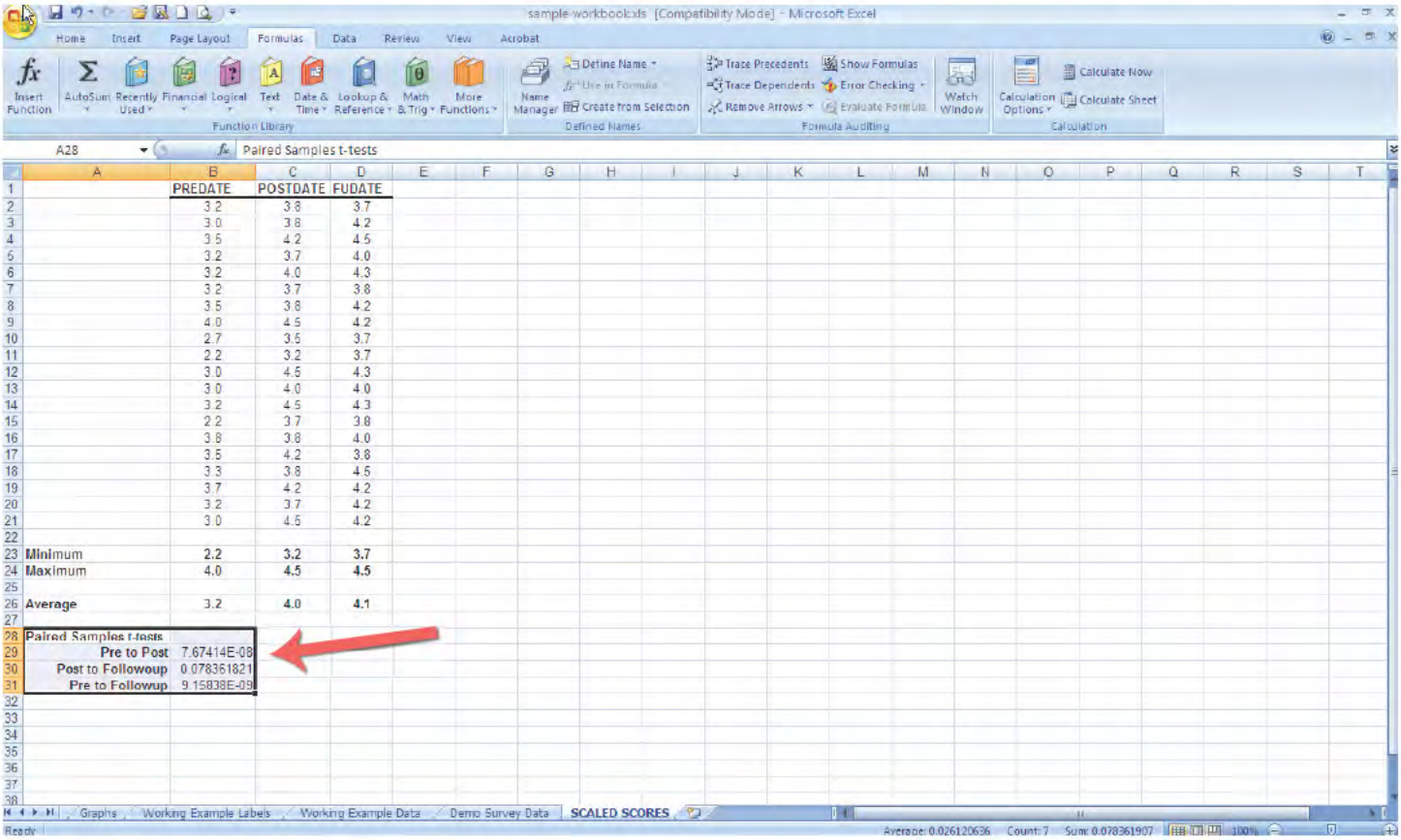


- After you click OK, a dialog box will open up. There will be four spaces in this dialog box you need to complete:
  - The **Array 1 box** tells you what the first set of data is that you want the function to be performed on. The cells are referred to by their column letter and row number. The colon in-between the cells means that Excel will look at all of the cells in between those two cells. You can change the cell range by typing in the cell range OR by using your mouse to highlight the cells you want included as shown in the picture below. **For a paired samples t-test the first array will be the scaled scores on the first survey you are interested in.**
  - The **Array 2 box** tells you what the second set of data is that you want the function to be performed on. It is completed the same as the Array 1 box. **For a paired samples t-test the second array will be the scaled scores on the second survey you are interested in.**
  - The **Tails** box should be filled in with the number **2**. This tells the computer that the difference could go two different ways: the first survey could be less than or greater than the second survey.
  - The **Type** box should be filled in with the number **1**. This tells the computer that the two arrays are **paired**. That is, the averages in each row are from the same person.

For example, for the CRCC data when the paired samples t-test to compare pre-survey data with post-survey data was run the dialog box looked like the picture below:

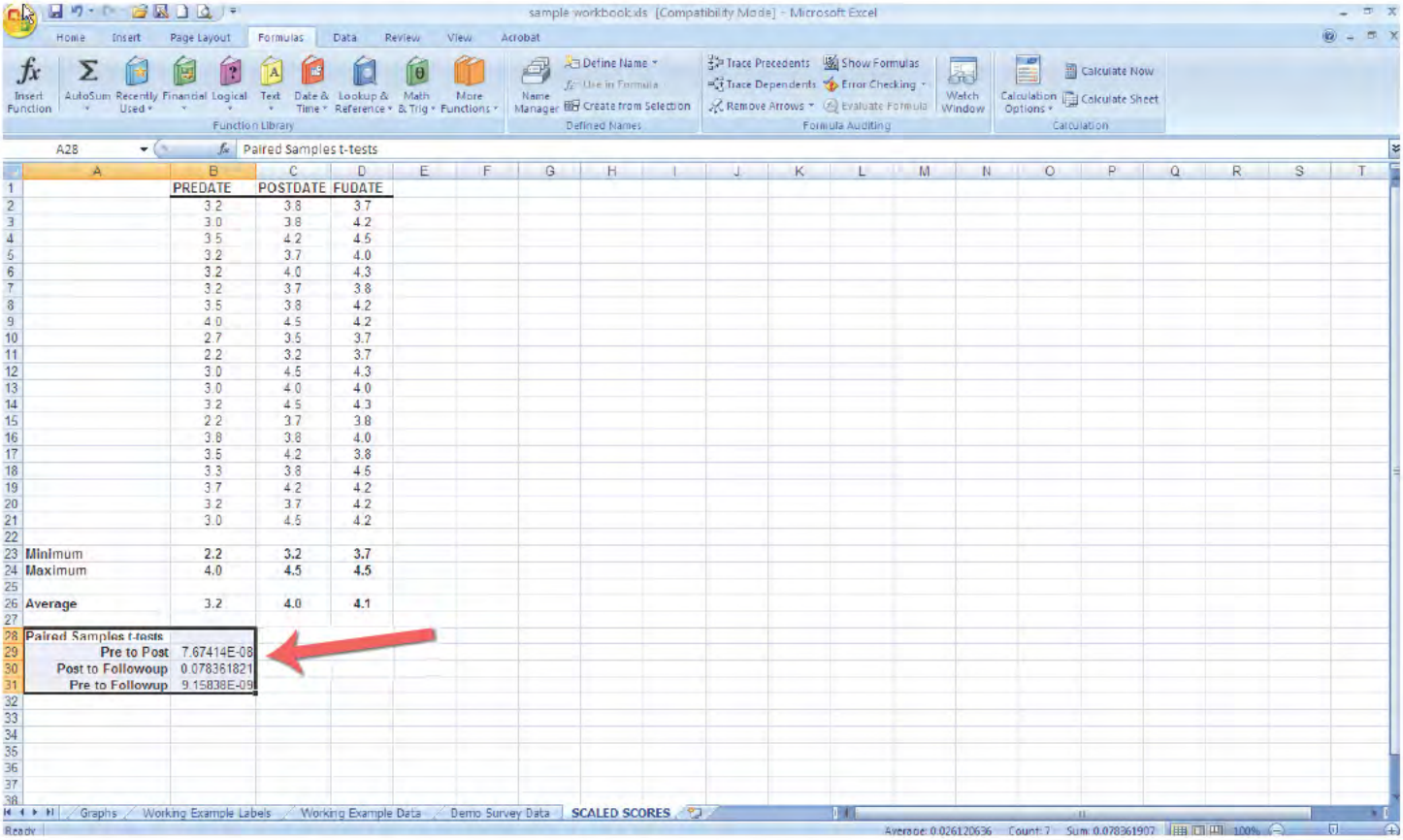


- Click OK. You will see that the test result for that comparison now appears.



- Repeat these steps to test the comparison for post-survey to follow-up survey and pre-survey to follow-up survey.

In the CRCC example, once all paired samples t-tests were completed they appeared as illustrated below:



	A	B	C	D
1		PREDATE	POSTDATE	FUDATE
2		3.2	3.8	3.7
3		3.0	3.8	4.2
4		3.5	4.2	4.5
5		3.2	3.7	4.0
6		3.2	4.0	4.3
7		3.2	3.7	3.8
8		3.5	3.8	4.2
9		4.0	4.5	4.2
10		2.7	3.5	3.7
11		2.2	3.2	3.7
12		3.0	4.5	4.3
13		3.0	4.0	4.0
14		3.2	4.5	4.3
15		2.2	3.7	3.8
16		3.8	3.8	4.0
17		3.5	4.2	3.8
18		3.3	3.8	4.5
19		3.7	4.2	4.2
20		3.2	3.7	4.2
21		3.0	4.5	4.2
22				
23	Minimum	2.2	3.2	3.7
24	Maximum	4.0	4.5	4.5
25				
26	Average	3.2	4.0	4.1
27				
28	Paired Samples t-tests			
29	Pre to Post	7.67414E-08		
30	Post to Followup	0.078361821		
31	Pre to Followup	9.15838E-09		
32				
33				
34				
35				
36				
37				
38				

For an explanation of how to interpret these results, see the next section on Interpreting Quantitative Data.

Independent Samples T-Test

The final test you are likely to want to run is an **independent samples t-test**. This test tells us whether two separate groups (i.e., "independent samples") give answers that are different from one another in a meaningful way. Like with the paired samples t-test, sometimes the differences between groups are very small. Therefore, we need to determine if the differences are meaningful or significant.

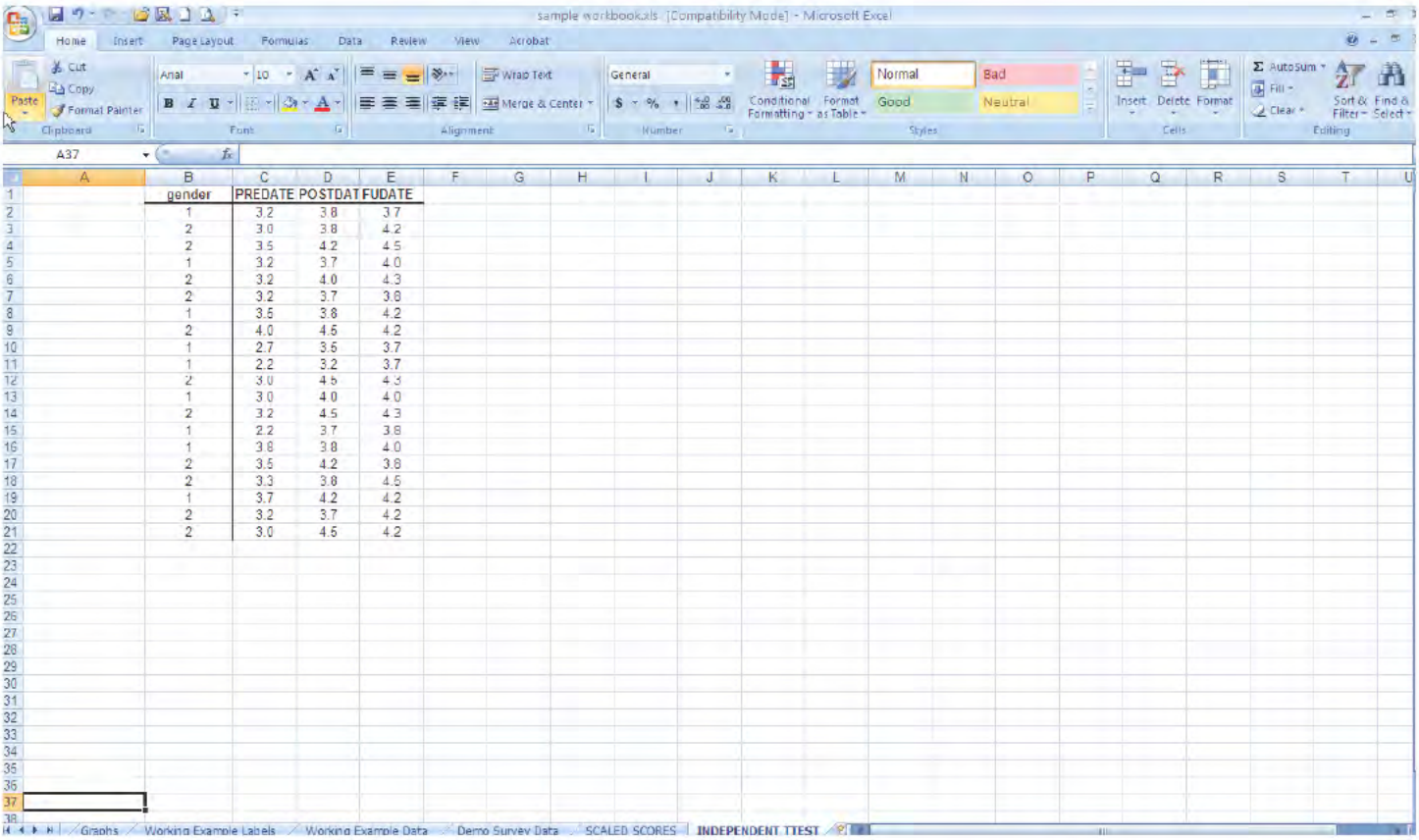
For example, CRCC staff may be interested to know if there are gender differences on any of the surveys. Therefore, they would want to compare the girls' pre-surveys and with boys' pre-surveys and then run similar comparisons for the post-surveys and follow-up surveys.

To run an independent samples t-test, Microsoft Excel requires the data first be sorted so that all of the scores for one group are together and all of the scores for the other group are together. Therefore, the first few steps for this test require that we organize our worksheet so the data are sorted by group.

Note: If you are comparing a group that participated in your program with a group that did not (e.g., a control or comparison group) then you would similarly use the independent samples t-test to analyze those comparisons. The procedures are the same as what will be illustrated here. The only difference is that you would be sorting people by which group (or "condition") they were in and comparing the average scores for each group.

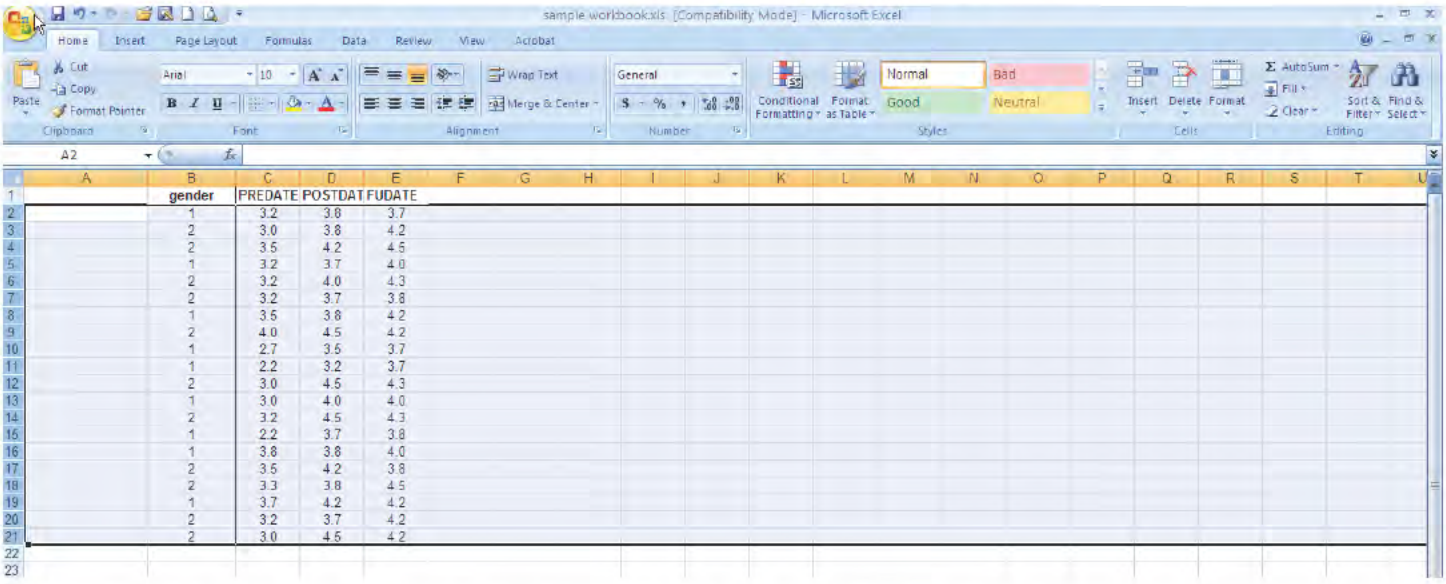
To run an independent samples t-test:

- From one of your previous worksheets, copy the column that designates which **group** each person belongs to and paste it below the results for the previous (paired samples) t-tests **OR** paste it into a new sheet.
- Then copy and paste your **scaled scores** next to the group data. ***Warning:*** **Make sure you are lining up each person's group with their scaled scores.** If you mix up the data you will get nonsensical results because the person's group (e.g., gender) will no longer be matched up with their survey scores.
- For example, in preparation for running the independent samples t-tests, the CRCC worksheet was formatted to look like the picture on the next page:



	gender	PREDATE	POSTDAT	FUDATE
2	1	3.2	3.8	3.7
3	2	3.0	3.8	4.2
4	2	3.5	4.2	4.5
5	1	3.2	3.7	4.0
6	2	3.2	4.0	4.3
7	2	3.2	3.7	3.8
8	1	3.5	3.8	4.2
9	2	4.0	4.5	4.2
10	1	2.7	3.5	3.7
11	1	2.2	3.2	3.7
12	2	3.0	4.5	4.3
13	1	3.0	4.0	4.0
14	2	3.2	4.5	4.3
15	1	2.2	3.7	3.8
16	1	3.8	3.8	4.0
17	2	3.5	4.2	3.8
18	2	3.3	3.8	4.5
19	1	3.7	4.2	4.2
20	2	3.2	3.7	4.2
21	2	3.0	4.5	4.2

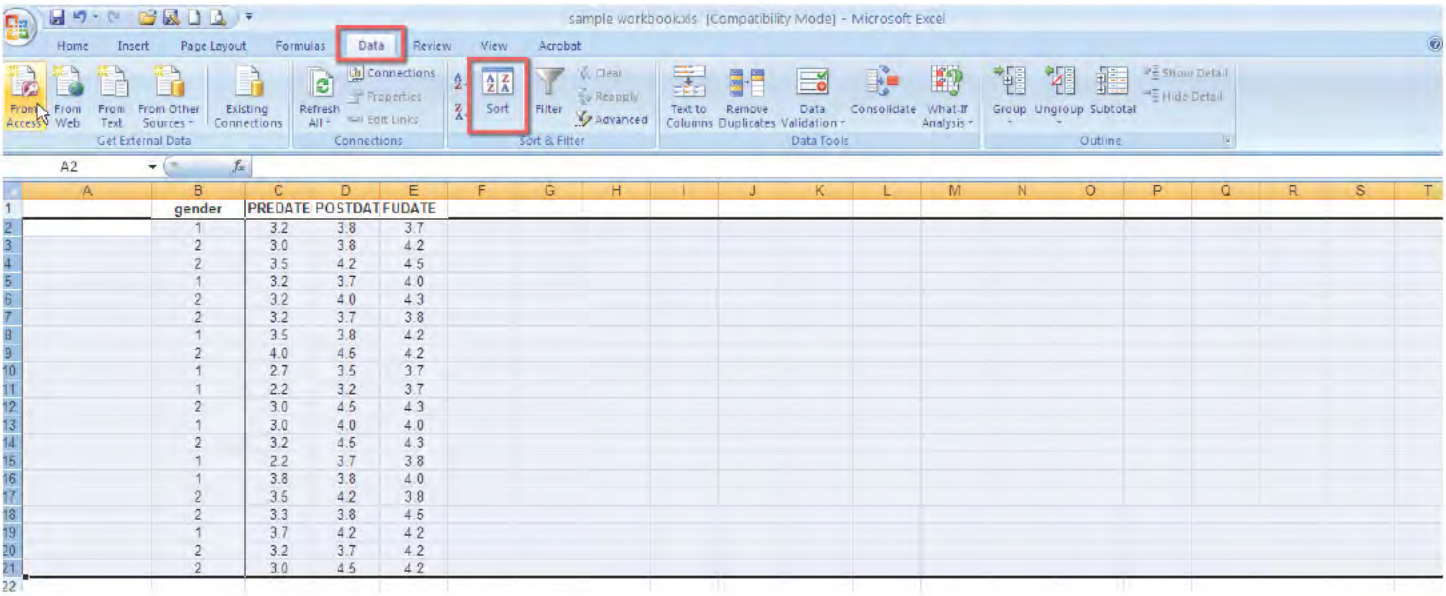
- Now you need to **sort** these rows so that all the boys' (gender = 1) scores are together and all the girls' (gender = 2) scores are together. *To do this:*
  - Click on the row numbers on the left side of the worksheet and drag your mouse down to highlight the **rows** containing the data you want to sort, as shown:



	gender	PREDATE	POSTDAT	FUDATE
2	1	3.2	3.8	3.7
3	2	3.0	3.8	4.2
4	2	3.5	4.2	4.5
5	1	3.2	3.7	4.0
6	2	3.2	4.0	4.3
7	2	3.2	3.7	3.8
8	1	3.5	3.8	4.2
9	2	4.0	4.5	4.2
10	1	2.7	3.5	3.7
11	1	2.2	3.2	3.7
12	2	3.0	4.5	4.3
13	1	3.0	4.0	4.0
14	2	3.2	4.5	4.3
15	1	2.2	3.7	3.8
16	1	3.8	3.8	4.0
17	2	3.5	4.2	3.8
18	2	3.3	3.8	4.5
19	1	3.7	4.2	4.2
20	2	3.2	3.7	4.2
21	2	3.0	4.5	4.2

Warning: It is critical that you highlight across the entire row. If you only highlight the "gender" column you will end up sorting the 1s and 2s, but the scaled scores will stay in the same order. This will mean that you have separated each person's gender from their scores.

- Click on the **Data ribbon** and then click the big **Sort button**. This will open up the Sort dialog box.

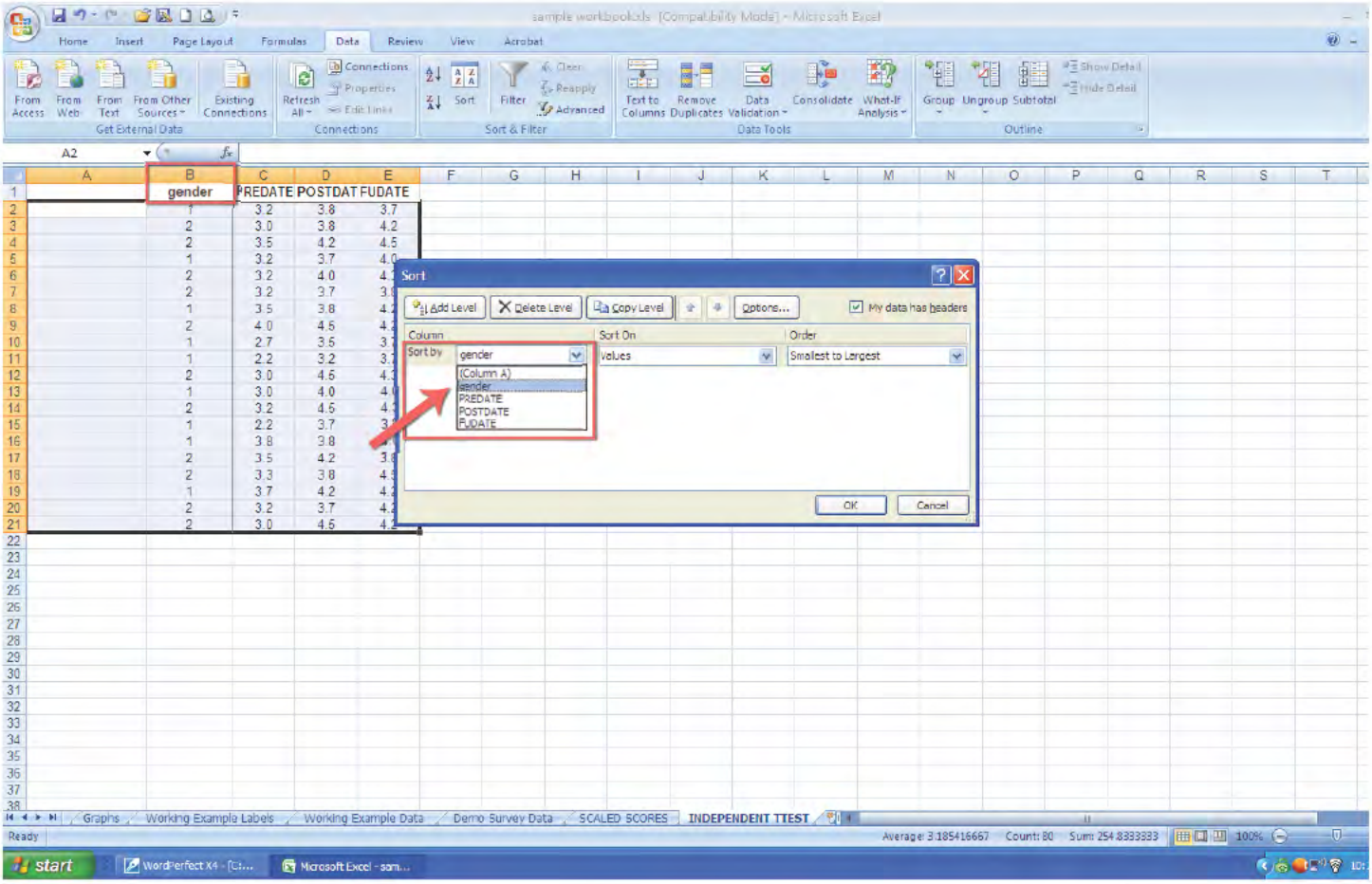


	gender	PREDATE	POSTDAT	FUDATE
2	1	3.2	3.8	3.7
3	2	3.0	3.8	4.2
4	2	3.5	4.2	4.5
5	1	3.2	3.7	4.0
6	2	3.2	4.0	4.3
7	2	3.2	3.7	3.8
8	1	3.5	3.8	4.2
9	2	4.0	4.5	4.2
10	1	2.7	3.5	3.7
11	1	2.2	3.2	3.7
12	2	3.0	4.5	4.3
13	1	3.0	4.0	4.0
14	2	3.2	4.5	4.3
15	1	2.2	3.7	3.8
16	1	3.8	3.8	4.0
17	2	3.5	4.2	3.8
18	2	3.3	3.8	4.5
19	1	3.7	4.2	4.2
20	2	3.2	3.7	4.2
21	2	3.0	4.5	4.2

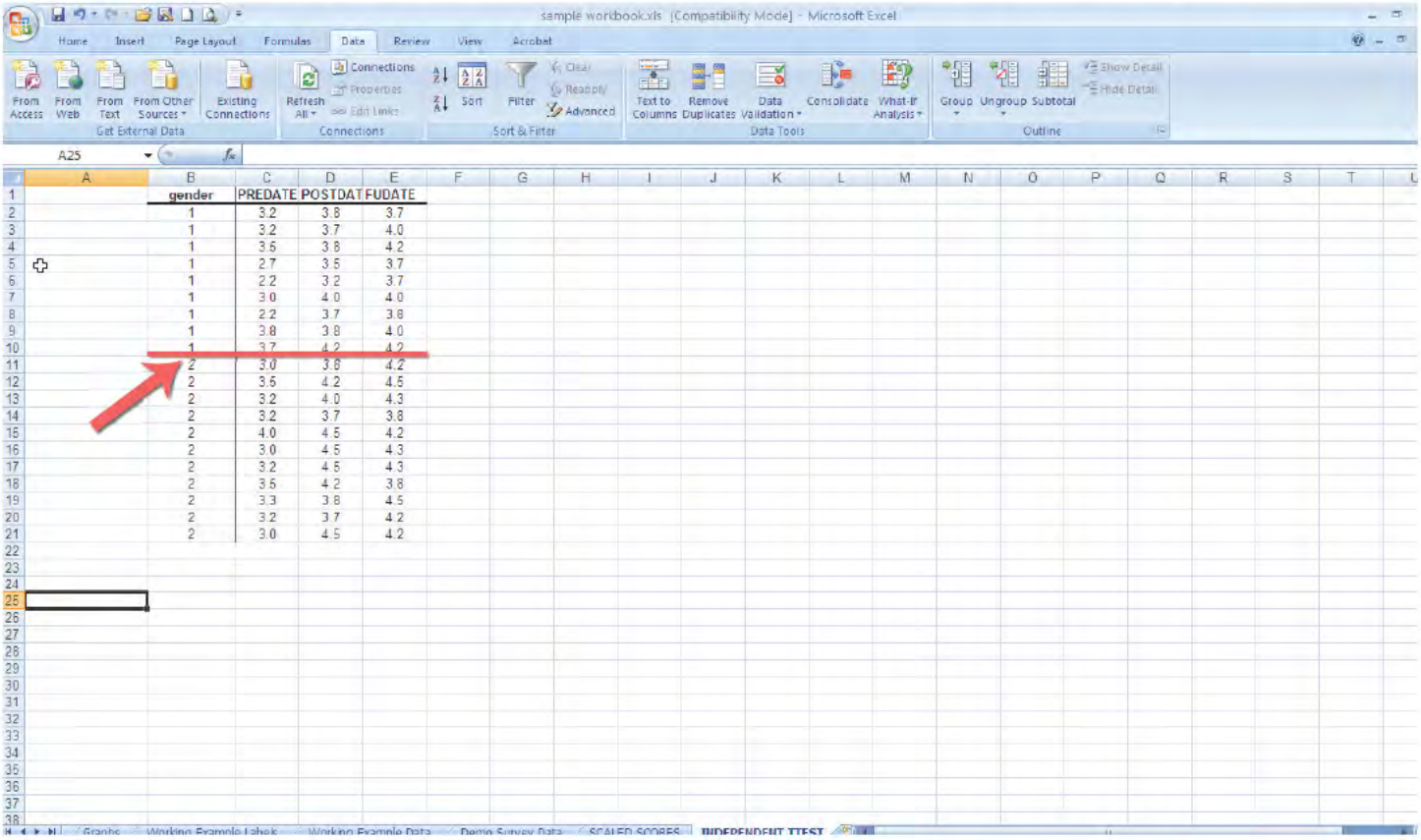
- In the Sort Dialog box, you need to **tell the program which column you want to sort by.** This will be the column that has the group information.

  In the CRCC example, the data will be sorted by gender (Column B) as shown below:

- Click OK.



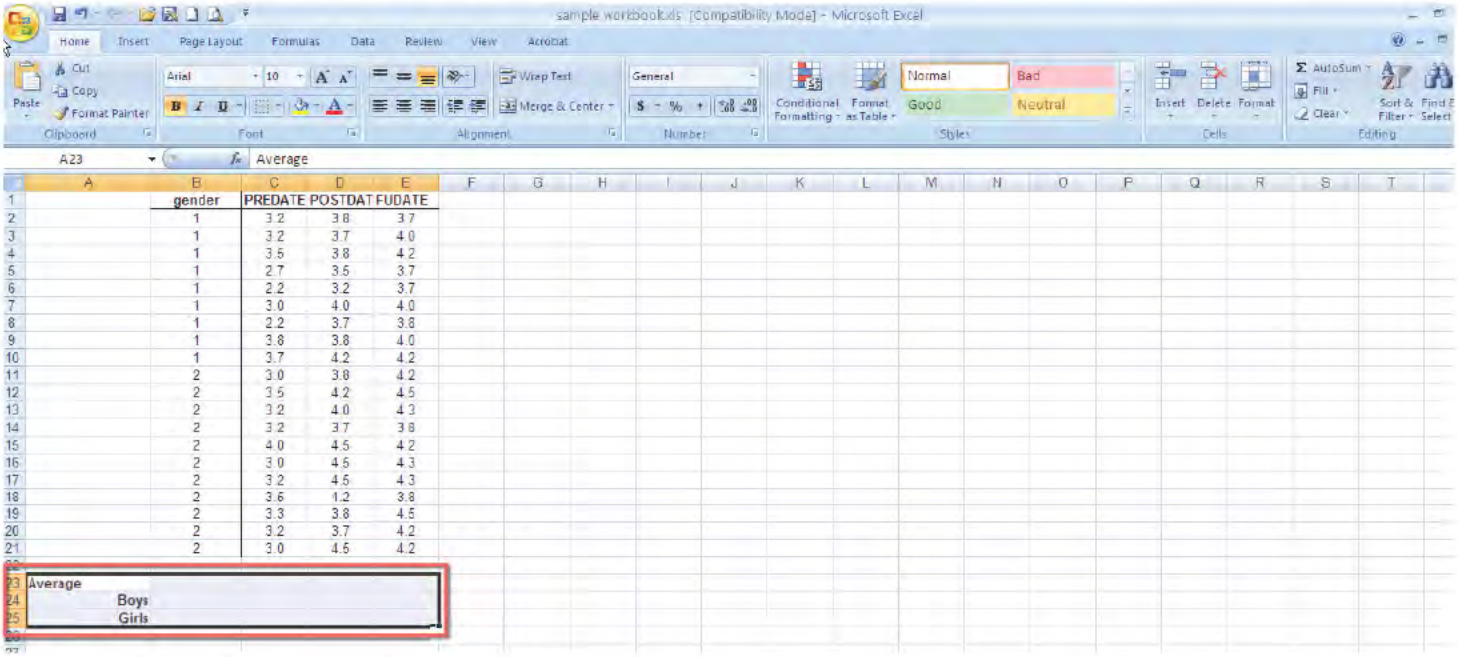
- **Now your data will be sorted by group,** as shown in the picture below. Notice that all the 1s (boys) are grouped together and all the 2s (girls) are grouped together. Because we highlighted across the entire row, when the 1s and 2s moved so did each person's scaled scores. So the data are still organized with one row representing one person's answers.



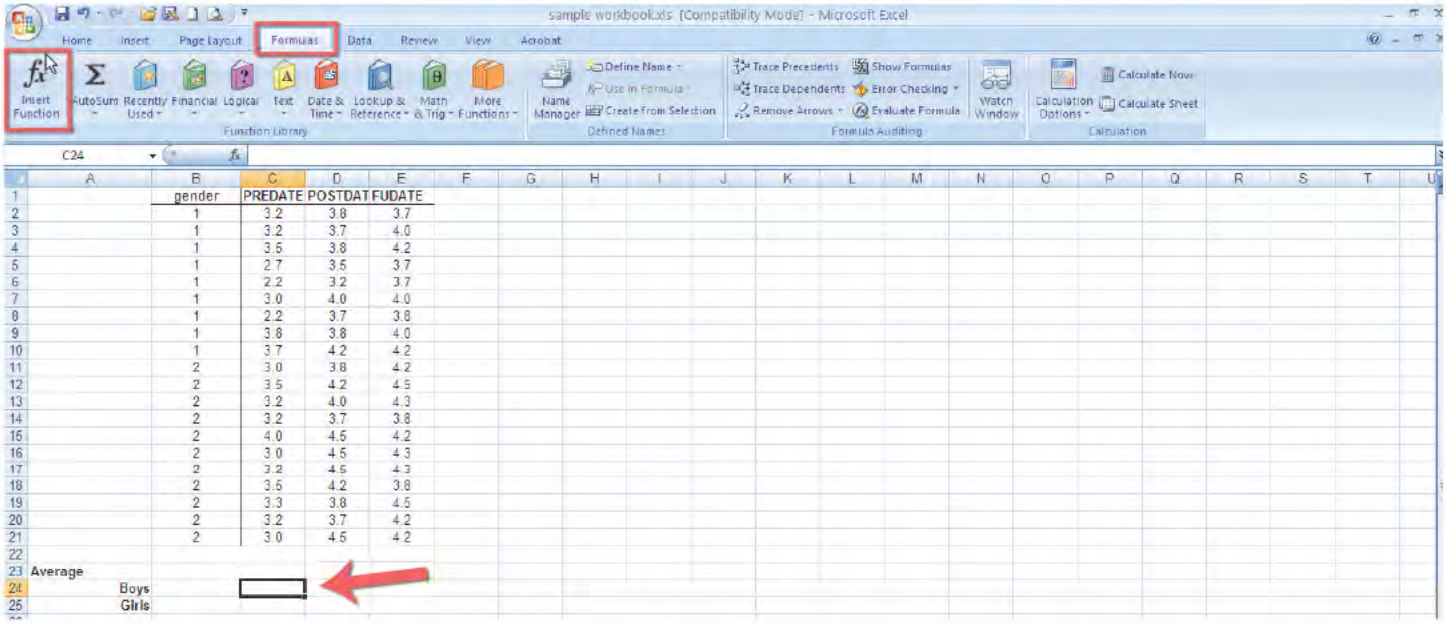
	gender	PREDATE	POSTDAT	FUDATE
2	1	3.2	3.8	3.7
3	1	3.2	3.7	4.0
4	1	3.5	3.8	4.2
5	1	2.7	3.5	3.7
6	1	2.2	3.2	3.7
7	1	3.0	4.0	4.0
8	1	2.2	3.7	3.8
9	1	3.8	3.8	4.0
10	1	3.7	4.2	4.2
11	2	3.0	3.8	4.2
12	2	3.5	4.2	4.5
13	2	3.2	4.0	4.3
14	2	3.2	3.7	3.8
15	2	4.0	4.5	4.2
16	2	3.0	4.5	4.3
17	2	3.2	4.5	4.3
18	2	3.5	4.2	3.8
19	2	3.3	3.8	4.5
20	2	3.2	3.7	4.2
21	2	3.0	4.5	4.2

Similar to the paired samples t-tests, the first look at group differences will be based on **average scores**. In this example, we are interested in what the **average score for boys** was on each survey versus the **average score for girls** on each survey. In your evaluation, you might be comparing groups other than by gender. Therefore, we need to calculate these group averages.

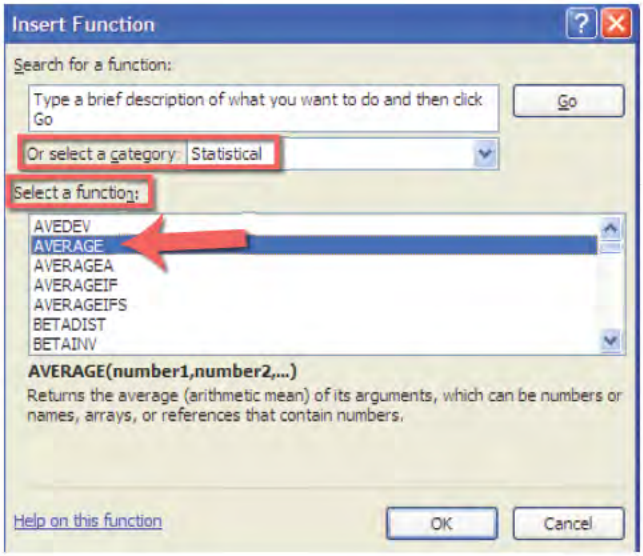
- Calculate group averages similar to how scale averages were calculated.
- Before you run the function, **label** where your averages will go. It is important to label your results from the beginning. The results will appear on the worksheet simply as numbers. Without labels you may quickly forget what the numbers refer to.



- **Click** to put your cursor in the cell where you want the first average to go. In the CRCC example, it would be in column B, row 24.
- In the **Formula ribbon**, click on the **Insert Function button**. This will open up the Function Dialog Box.

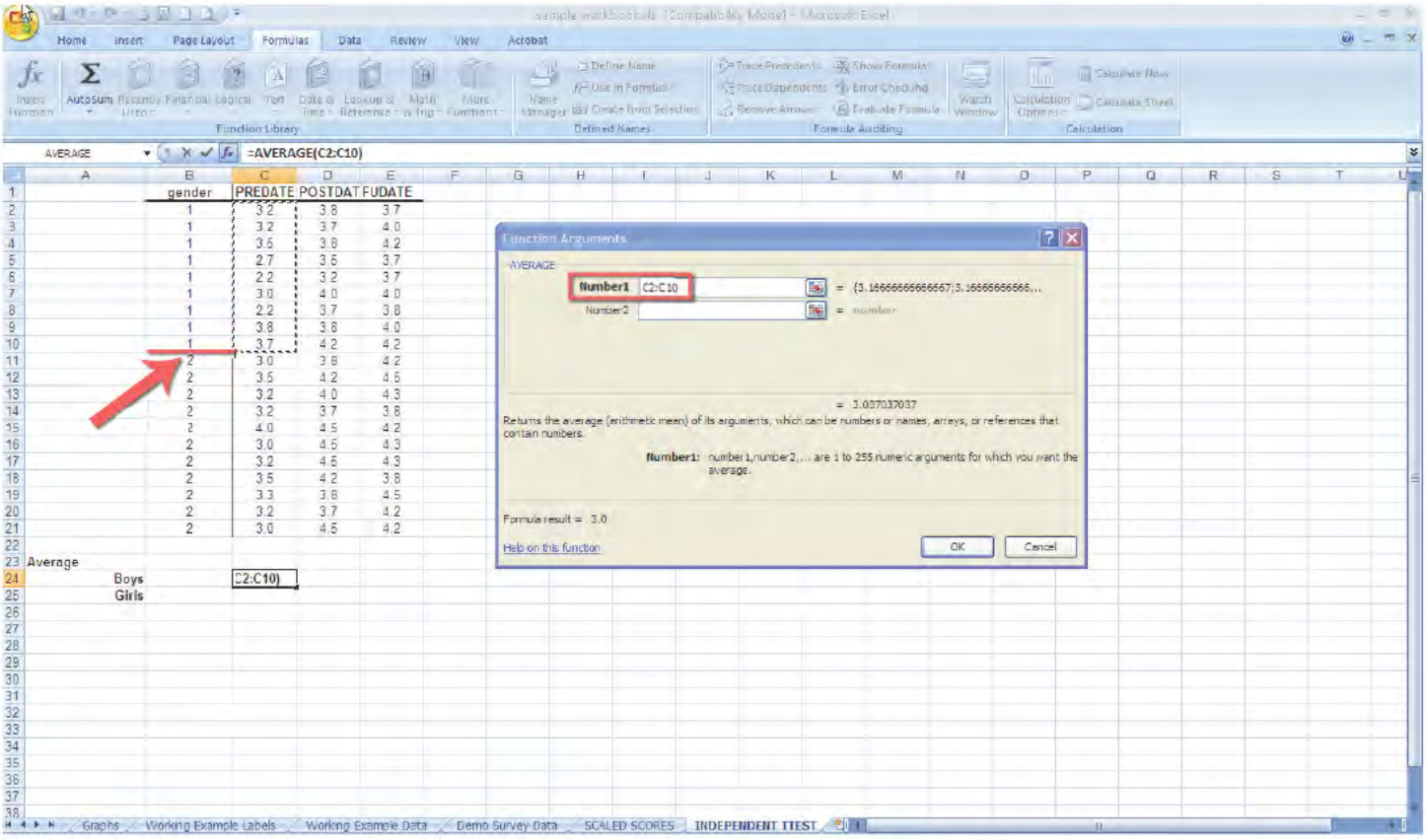


- In the Function Dialog Box make sure that where it says "Or select a category" the category is "**Statistical.**" If it is not, then use the drop-down list to select the Statistical functions.
- In the "Select a function" box scroll until you find the **AVERAGE** function. Highlight AVERAGE and click OK. ***Note***: There are multiple functions for averaging. Make sure you are using the plain AVERAGE function. Do not use AVEDEV or AVERAGE A. These are different functions and will give you different results.



- After you click OK, a dialog box will open up. The **range box** tells you what cells the function will be performed on. The cells are referred to by their column letter and row number. The colon in-between the cells means that Excel will look at all of the cells in between those two cells. You can change the cell range by typing in the cell range OR by using your mouse to highlight the cells you want included as shown on the next page.

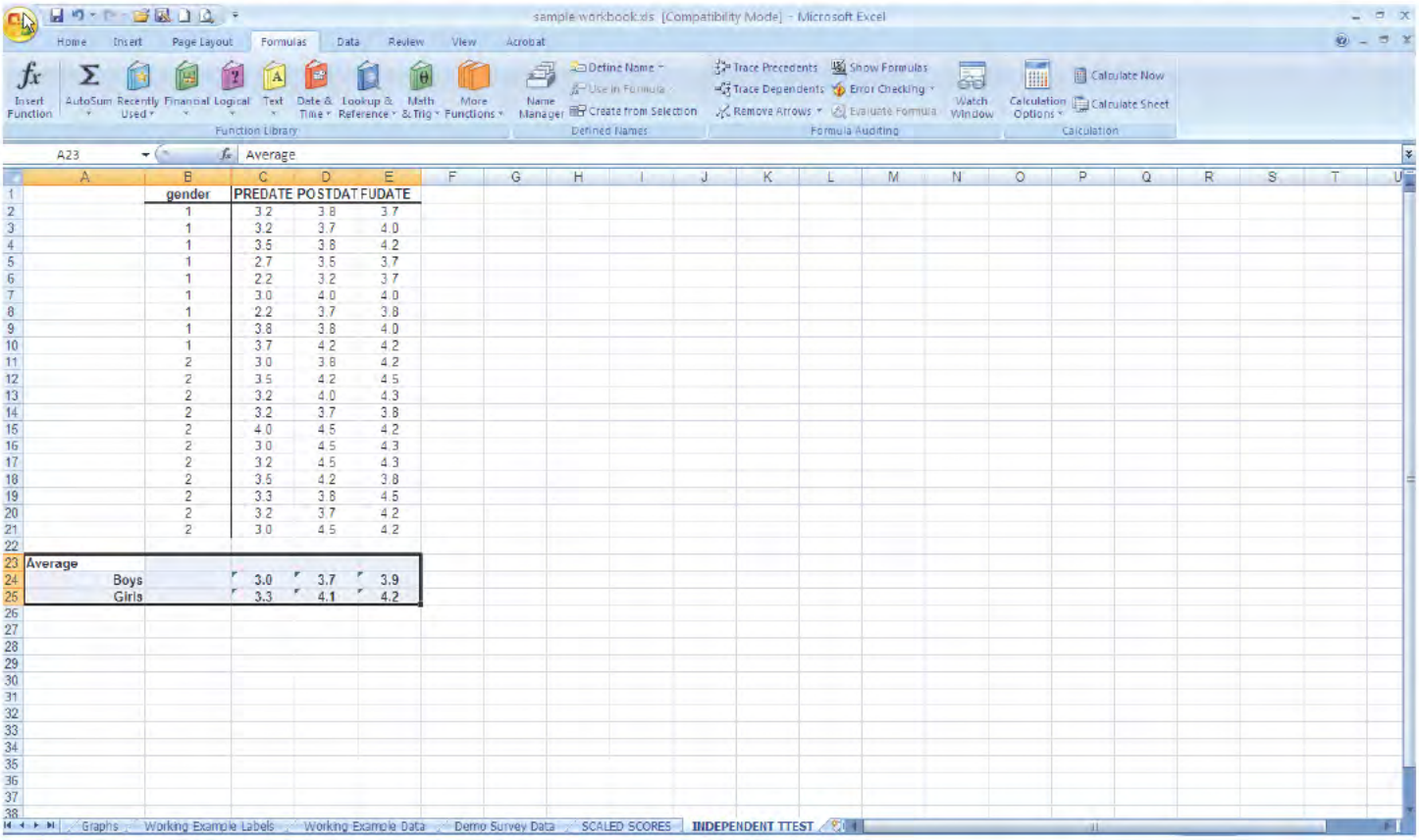
  Warning: **Make sure when you designate the range you only include one group's data.** For example, if you are comparing by gender, calculate the average for boys first. Then you will repeat these steps to calculate the average for girls.



  - **Click OK.** You will see that the average for that group on that test now appears.

- Once you have calculated the first average you can automatically repeat that function for that group all of the remaining scales. You can do this by holding your mouse so the cursor hovers over the lower right corner of that cell where you have the first average. You will see that a + sign appears. Hold down on the mouse and drag that + sign across the row, highlighting all of the cells as you go. When you let go, the averages for each scale will appear in the appropriate places.
- **Repeat** these steps to calculate the averages for the other group.

In the CRCC example, once all average scores were completed they appeared as illustrated below:



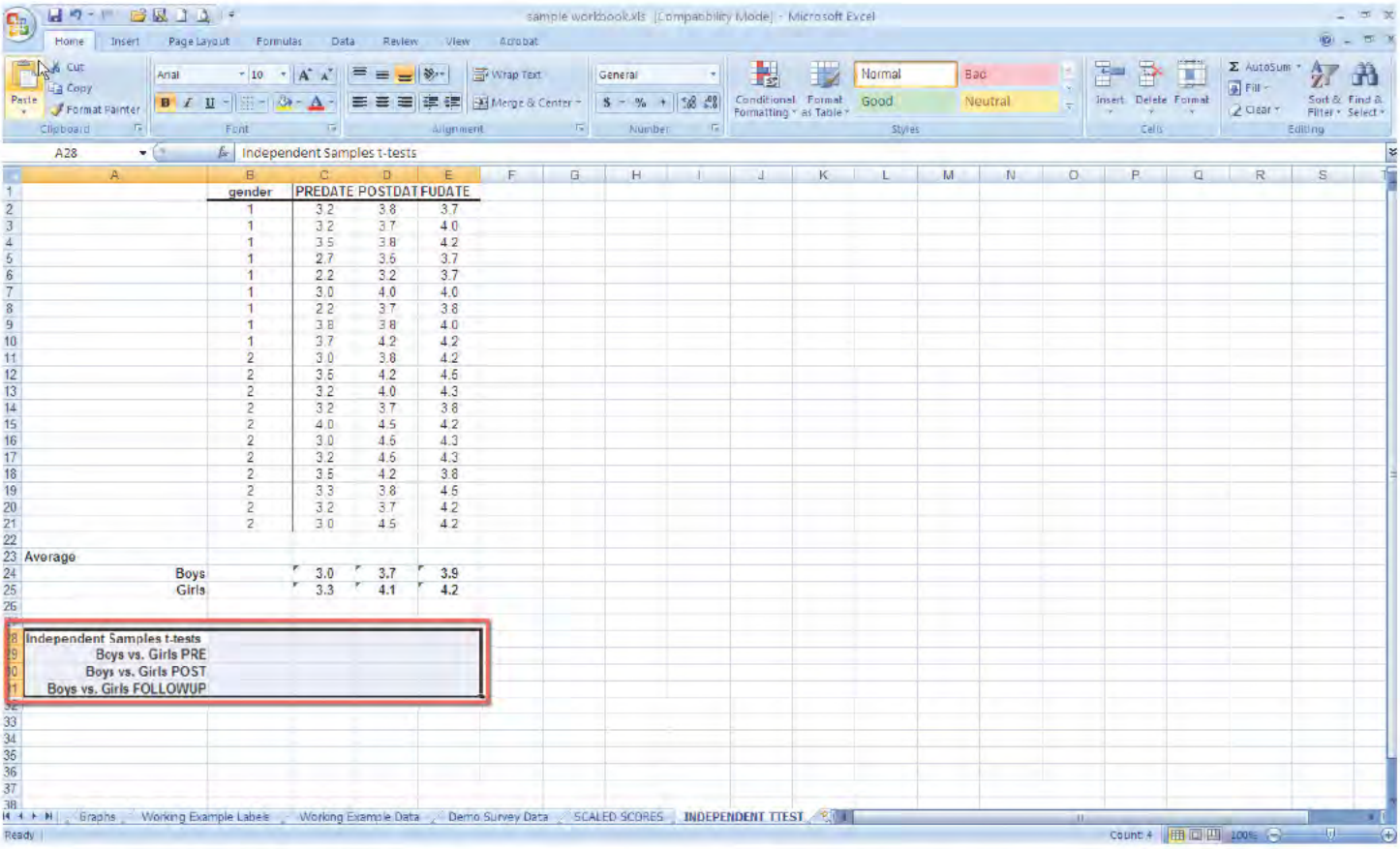
	A	B	C	D	E
1		gender	PREDATE	POSTDAT	FUDATE
2		1	3.2	3.8	3.7
3		1	3.2	3.7	4.0
4		1	3.5	3.8	4.2
5		1	2.7	3.5	3.7
6		1	2.2	3.2	3.7
7		1	3.0	4.0	4.0
8		1	2.2	3.7	3.8
9		1	3.8	3.8	4.0
10		1	3.7	4.2	4.2
11		2	3.0	3.8	4.2
12		2	3.5	4.2	4.5
13		2	3.2	4.0	4.3
14		2	3.2	3.7	3.8
15		2	4.0	4.5	4.2
16		2	3.0	4.5	4.3
17		2	3.2	4.5	4.3
18		2	3.5	4.2	3.8
19		2	3.3	3.8	4.5
20		2	3.2	3.7	4.2
21		2	3.0	4.5	4.2
22					
23	Average				
24	Boys		3.0	3.7	3.9
25	Girls		3.3	4.1	4.2

We are now ready to run the independent samples t-tests.

Before you run the function, **label where your test results will go.** It is important to label your results from the beginning. The results will appear on the worksheet simply as numbers. Without labels you may quickly forget what the numbers refer to.

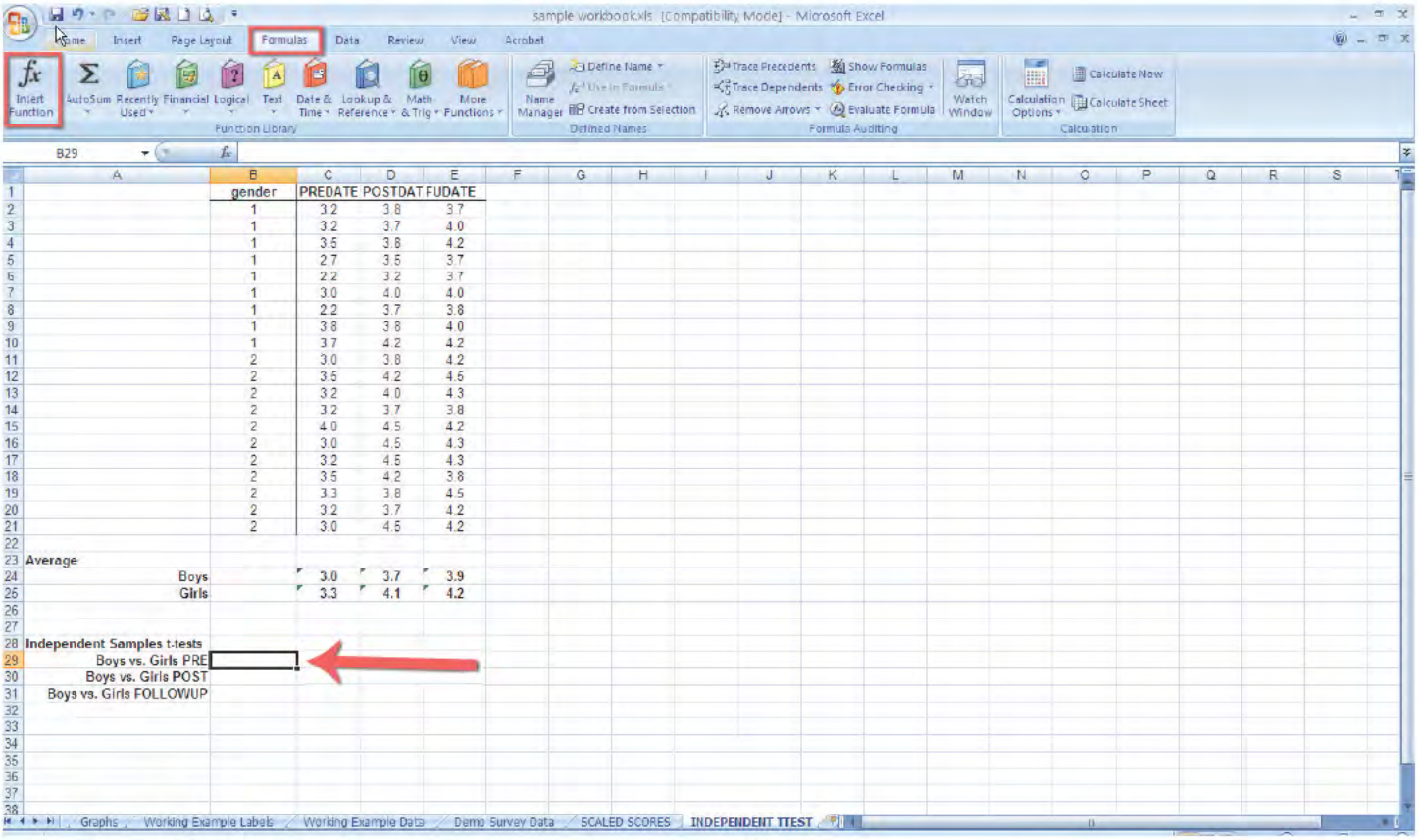
- As shown in the picture below, it is recommended that you put t-test results below the averages.
- The number of independent samples t-tests you run will depend on how many comparisons you are making. Each comparison is one test. For example, CRCC has three comparisons to make:

  1. Boys vs. Girls for Pre-survey
  2. Boys vs. Girls for Post-survey
  3. Boys vs. Girls for Follow-up survey

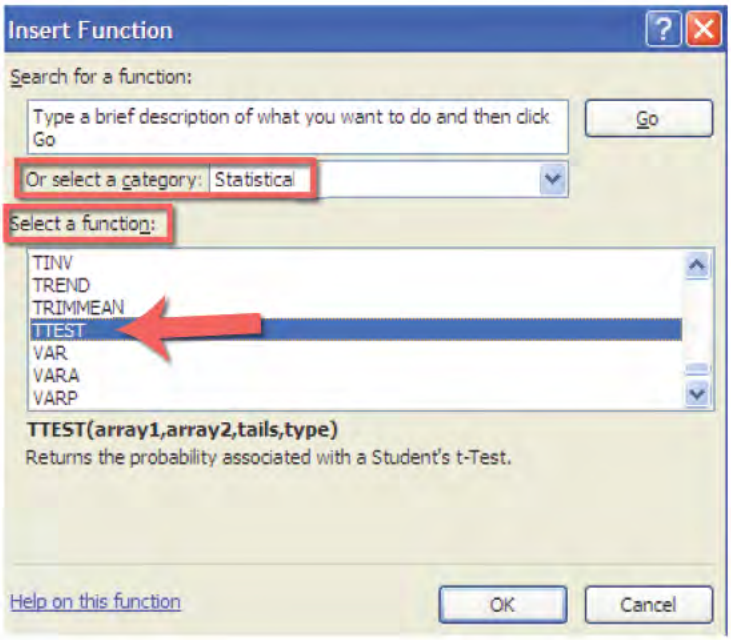


	gender	PREDATE	POSTDAT	FUDATE
	1	3.2	3.8	3.7
	1	3.2	3.7	4.0
	1	3.5	3.8	4.2
	1	2.7	3.5	3.7
	1	2.2	3.2	3.7
	1	3.0	4.0	4.0
	1	2.2	3.7	3.8
	1	3.8	3.8	4.0
	1	3.7	4.2	4.2
	2	3.0	3.8	4.2
	2	3.5	4.2	4.5
	2	3.2	4.0	4.3
	2	3.2	3.7	3.8
	2	4.0	4.5	4.2
	2	3.0	4.5	4.3
	2	3.2	4.5	4.3
	2	3.5	4.2	3.8
	2	3.3	3.8	4.5
	2	3.2	3.7	4.2
	2	3.0	4.5	4.2
Average				
Boys		3.0	3.7	3.9
Girls		3.3	4.1	4.2

- To calculate the **independent samples t-test** results, you will use Excel's functions much like you did when you calculated the paired samples t-test. You will be using the TTEST function.
- **Click** to put your cursor in the cell where you want the first t-test result to go. In the CRCC example, it would be in column C, row 28.
  - In the **Formula ribbon**, click on the **Insert Function button**. This will open up the Function Dialog Box.

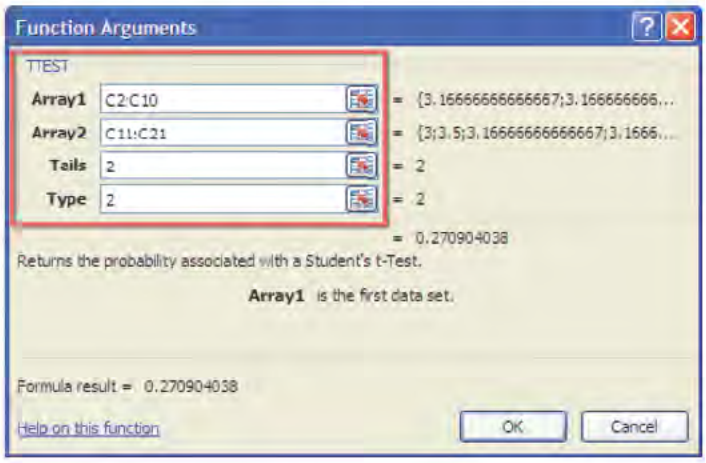


- In the Function Dialog Box make sure that where it says "Or select a category" the category is "**Statistical.**" If it is not, use the drop-down list to select the Statistical functions.
- In the "Select a function" box scroll until you find the TTEST function. Highlight **TTEST** and click **OK.**



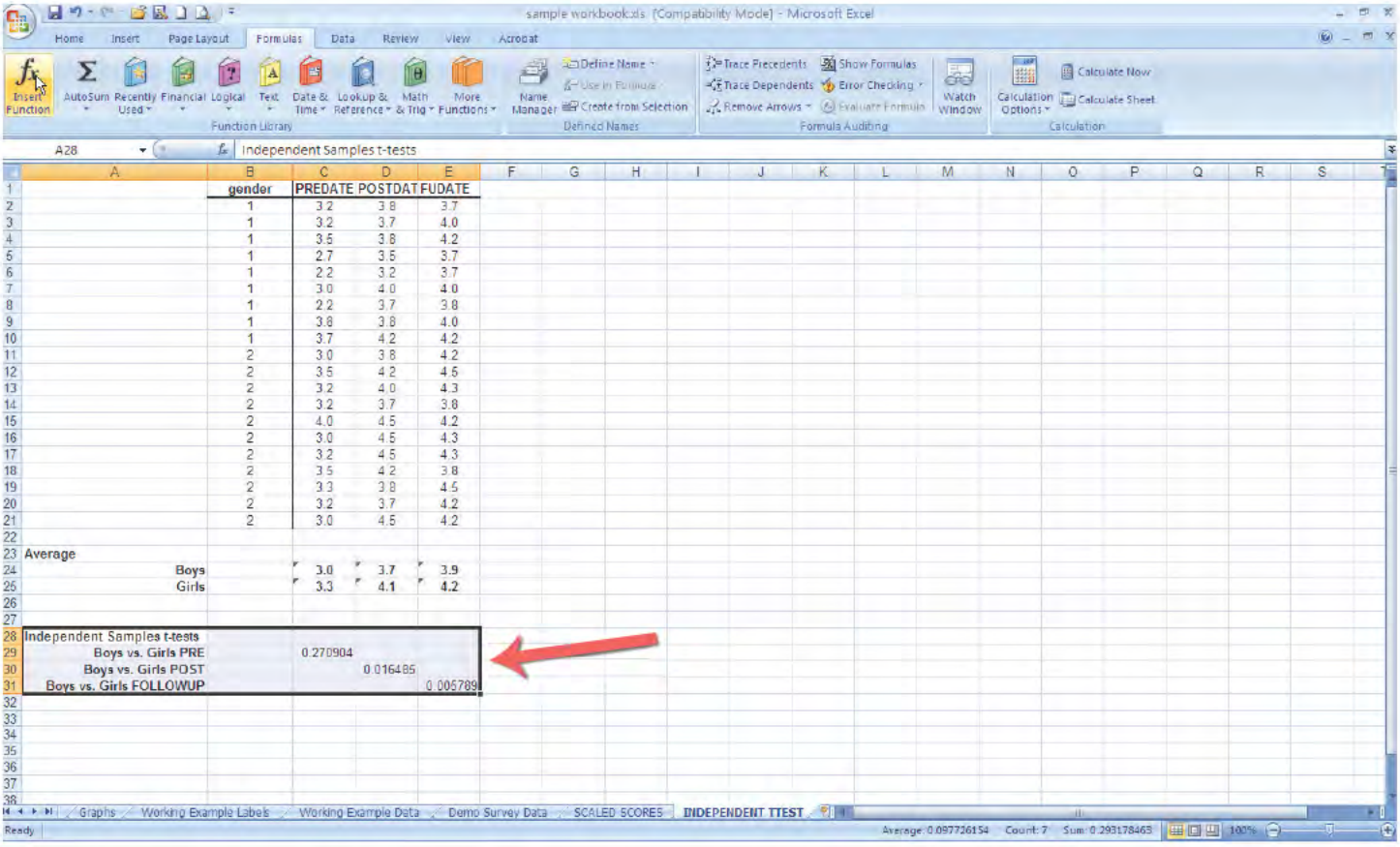
- After you click OK, a dialog box will open up. There will be four spaces in this dialog box you need to complete:
  - The **Array 1 box** tells you what the first set of data is that you want the function to be performed on. The cells are referred to by their column letter and row number. The colon between the cells means that Excel will look at all of the cells in between those two cells. You can change the cell range by typing in the cell range or by using your mouse to highlight the cells you want included as shown in the picture below. **For an independent samples t-test the first array will be the first group in which you are interested.**
  - The **Array 2 box** tells you what the second set of data is on which you want the function to be performed. It is completed the same as the Array 1 box. **For an independent samples t-test the second array will be second group you are interested in.**
  - The **Tails** box should be filled in with the number **2**. This tells the computer that the difference could go two different ways: the first group's scores could be less than or greater than the second group's scores.
  - The **Type** box should be filled in with the number **2**. This tells the computer that the two arrays are **independent**. That is, they are from two different groups.

    For example, for the CRCC data when the independent samples t-test to compare by gender was run the dialog box looked like the picture below:



- **Click OK.** You will see that the test result for that comparison now appears.
- **Repeat** these steps to test the comparison for gender differences on the post-survey and follow-up survey data.

In the CRCC example, once all independent samples t-tests were completed they appeared as illustrated below:



	gender	PREDATE	POSTDAT	FUDATE
	1	3.2	3.8	3.7
	1	3.2	3.7	4.0
	1	3.5	3.8	4.2
	1	2.7	3.5	3.7
	1	2.2	3.2	3.7
	1	3.0	4.0	4.0
	1	2.2	3.7	3.8
	1	3.8	3.8	4.0
	1	3.7	4.2	4.2
	2	3.0	3.8	4.2
	2	3.5	4.2	4.5
	2	3.2	4.0	4.3
	2	3.2	3.7	3.8
	2	4.0	4.5	4.2
	2	3.0	4.5	4.3
	2	3.2	4.5	4.3
	2	3.5	4.2	3.8
	2	3.3	3.8	4.5
	2	3.2	3.7	4.2
	2	3.0	4.5	4.2
Average				
Boys		3.0	3.7	3.9
Girls		3.3	4.1	4.2
Independent Samples t-tests				
Boys vs. Girls PRE		0.270904		
Boys vs. Girls POST			0.016485	
Boys vs. Girls FOLLOWUP				0.005789

The analyses covered in this manual are now complete and we can turn to the question of how to interpret these results.

CHECKLIST

In this section you learned how to run four basic analyses.

Task 1: Create Scaled Scores (If necessary)

- Identify the questions that make up your scale.
- Decide if the scale will be an average or a sum.
- Insert a column in your worksheet for the scaled score and label it.
- Use the Function tool to create your scaled score for the first person.
- Copy the function to create scaled scores for all of the remaining people.

Task 2: Identify the Types of Date You Have

- Decide what you want to know.
- Identify which columns (variables) have the relevant data.
- Identify the type of data those variables are: Categorical or Interval.

Task 3: Identify the Appropriate Analyses

- Use the table on Page 63 of this manual to determine which tests you will need to run to answer your evaluation questions.

Task 4: Run the Analyses

- Calculate Frequencies.
- Calculate Percentages.
- Calculate Averages for Scaled Scores.
- Calculate Paired Samples T-Tests to test for change over time.
- Sort your data into groups you want to compare.
- Calculate Averages for the groups you want to compare.
- Calculate Independent Samples T-Tests to test for differences between groups.

INTERPRETING QUANTITATIVE EVALUATION DATA

- Task 1: Frequencies and Percentages
- Task 2: Averages — Changes Over Time
- Task 3: Paired Samples T-Tests
- Task 4: Averages — Differences Between Groups
- Task 5: Independent Samples T-Tests
- Task 6: Decide if Follow-up Analyses are Needed
- Task 7: Summarize and Report Your Findings

INTERPRETING QUANTITATIVE RESULTS

After running the appropriate analyses, the next step is to look at the results and interpret what they mean. To make this task less daunting, think of it as telling a story. The numbers are telling us a story. We need to figure out what that story is. As with any event, how we tell the story will be determined by many factors including what we want to know, what stood out to us and how one train of thought leads to another. The suggestions and examples provided in this manual are intended as illustrations to stimulate your thinking. If what you want to know is different or if your data looks very different, you may need to take a different approach to interpreting your results. However, the examples provided here reflect some of the most common evaluation questions program staff often have, so you will likely find they are applicable to your own data.

Let the data tell a story.

To help you organize your thoughts and focus your attention, it is helpful to write up your results and interpretations on a separate sheet of paper (or in a word processing file if you prefer to write on the computer).

This section will follow, in order, the types of analyses that you learned about in the previous section of the manual. Each set of guidelines for interpreting results will be followed by looking at the results from the CRCC example and interpreting those results.

TASK 1: FREQUENCIES AND PERCENTAGES

Frequencies and percentages tell us **how often** a particular answer was given, either in raw numbers (frequencies) or proportions (percentages). These two results are related because frequencies are used when calculating percentages. Therefore, they often tell us the same thing.

As mentioned earlier, frequencies can be compared only when the scales they are based on are the same (e.g., all answers were given on the same 1–5 scale) or if you are comparing groups when there are the same number of people in each group. Percentages can be compared even when you have different scales. If the scales are the same, the interpretation of the frequencies and percentages will be identical. **Most people find it easier to think in terms of percentages, so you might want to use only the percentages in your interpretations.**

Both frequency and percentage results are useful for thinking about answers to individual questions. What you are looking for are patterns. Some questions to consider include:

1. Were any answers never used?
2. What answers were used most/least?
3. Were there different patterns depending on the question?

CASE EXAMPLE

CRCC Percentages

CRCC staff were interested in how students responded on each of the surveys. The first thing they were curious about was how often each answer was given. They expected that some students would rely heavily on gender stereotypes, some students would be in the middle and some students would be very strong in their endorsement of gender equity. If this was the case, they would expect to see each answer used about the same number of times.

The percentage results for the pre-survey were as follows:

	PAY FOR DATE (recoded)	ASK FOR DATE	SEX EQUITY	LOOK SMART (recoded)	HOLD DOOR, ETC.
1 (Strongly Disagree)	5%	**25%**	15%	0%	10%
2 (Mostly Disagree)	20%	**30%**	30%	10%	30%
3 (Slightly Disagree)	30%	**20%**	35%	10%	25%
4 (Slightly Agree)	20%	10%	5%	**25%**	20%
5 (Mostly Agree)	10%	10%	10%	**25%**	10%
6 (Strongly Agree)	10%	5%	5%	**20%**	5%

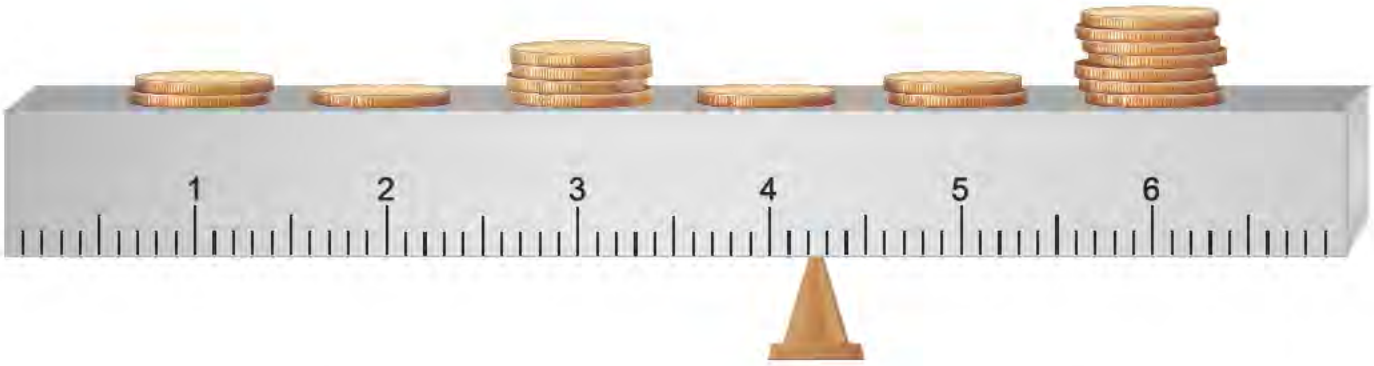
Remember: For all of these questions higher numbers indicate greater endorsement of gender equity.

Interpretations: The patterns that stood out to CRCC staff most were:

- For most questions every answer was used. This indicates that there was a wide range in students' attitudes.
- How much students endorsed gender equity depended on the behavior being asked. For example, looking smarter than your date showed a lot of gender equity but who asks for a date was more likely to be looked at based on gender stereotypes.

TASK 2: AVERAGES: CHANGES OVER TIME

Averages provide us a summary of where the scores on a scale gather round. You might imagine this like putting pennies on a ruler. In the case of the CRCC survey, there was a six-point scale, so imagine a ruler. On that ruler, put one penny to represent each person's scaled score for attitudes toward dating and heterosexual relationships. For example, if someone had a scaled score on the pre-survey of 2.0, their penny would go on the two-inch mark. If they had a scaled score of 4.0, their penny would go on the four-inch mark. People who have the same score will have their pennies stacked on top of one another. So you end up with a ruler that might look something like this:



If we tried to balance this ruler, where would we need to put the **balancing point**? That location represents the average. In this illustration, because most of the pennies are either at the mid-point or higher, the average will be somewhere between 3 and 6. (The average is 4.3.)

It is important to remember that an average score does not mean that every person's answer was near the average. In this example, we see that three pennies are much lower than the average. By definition, there will always be some people who responded above the average and some people who responded below the average.

Consequently, you need to be somewhat cautious when relying on averages. For example, the "average" annual temperature in Williamsport, PA, is 49.9°F. If you were coming from Helsinki, Finland to Williamsport on vacation in July and that is all the information you had when you were packing, you might end up having to buy a lot of summer clothes when you arrived![11]

11 Those who have a background in statistics may recall that this is why whenever we report means we also report additional statistics such as standard deviations and why we screen for outliers before running parametric tests. However, for the sake of people who do not have any prior training in statistics, in this manual we will keep things simple.

Despite this limitation, averages are one of the most useful ways of summarizing our data. They provide a simple way of looking at how scores change over time. Specifically, in the evaluation designs most programs use for their prevention programs (as well as for counseling services) the key questions we are interested in are:

- How did scores change from the pre-survey to the post-survey?
- How did scores change from the post-survey to follow-up survey?
- If there was no change from pre-survey to post-survey, how did scores change from the pre-survey to the follow-up survey?

What we are hoping to see is:

- Positive changes from pre-survey to post-survey
- No negative change (e.g., reverting back to old attitudes) from post-survey to follow-up survey.
- If there was no change from pre-survey to post-survey, then we hope to see positive changes from pre-survey to follow-up survey. Perhaps the changes just took longer to happen.
- If there was a negative change from post-survey to follow-up survey, then we hope that the follow-up numbers are at least higher than the pre-survey. They may have reverted back to old attitudes a little but didn't lose all the effects of the program.

CASE EXAMPLE

CRCC Averages on Scaled Scores

CRCC staff were hoping that their program would increase students' endorsement of gender equity in dating and heterosexual relationships. This would be reflected by their scaled scores increasing over time.

The average scaled scores were as follows (see Page 66 for the Excel worksheet where these results came from):

Pre-survey average	**= 3.2**
Post-survey average	**= 4.0**
Follow-up survey average	**= 4.1**

Remember: For all of these questions higher numbers indicate greater endorsement of gender equity.

Interpretations: What CRCC staff saw:

- Scores increased over time with each survey.
- The change from pre-survey to post-survey was quite noticeable.
- But the change from post-survey to follow-up survey was very small.

TASK 3: PAIRED SAMPLES T-TEST

Any differences between the surveys raise the question: is this difference meaningful? In statistics we ask the question: is this difference **statistically significant**? The precise meaning of statistical significance is beyond the scope of this manual.[12] Instead, we will simplify it to the idea: is this difference "real" or is it simply due to random fluctuations?

It is worth our time to consider what it means for a difference to be due to **random** fluctuations. Let us use another conceptual example to illustrate this. Imagine that today you took the written test to obtain a Pennsylvania driver's license. You scored 90 percent correct. Six weeks later you take the test again and score 85 percent correct. Are you really a worse driver than you were six weeks earlier? Or is the difference in your scores due to random fluctuations such as: you were more interested the first time, you were more tired the second time, the second time you took the test you were distracted by thinking about a crisis at work, etc.? Chances are your lower score was due to random fluctuations and not due to your really being a worse driver than you were six weeks ago.

Similarly, people who participate in our programs or who receive services from our agencies will not always give the same answers every time on a survey. They will be influenced by fatigue, distractions, motivation, noise, etc. What we need to determine is whether or not we think the differences in scores are big enough to justify saying that there was a real change in their attitudes, knowledge, behaviors, etc.

To make this judgment, we calculate something called a **p-value** (*p* stands for *probability*).[13] This is the number that appeared in your worksheet after you ran the paired samples t-test. You have a separate p-value for each comparison you made. In the CRCC example, there were three p-values calculated: (1) pre- to post-survey, (2) post- to follow-up survey and (3) pre- to follow-up survey.

- If the p-value is **less than 0.05**, we say that the difference is **statistically significant**. In other words, there really was a difference.
  - After determining that there is a significant difference we can look at the **averages** to see whether the change was due to scores increasing or decreasing. Do not assume that a significant difference means you got the results you wanted! Always go back and check the averages.
- If the p-value is **0.05 or greater**, we say that the difference is **not** statistically significant. In other words, there was no real difference. Any difference was due to random fluctuations and not due to real changes.

12 For those who are curious, t-tests statistical significance is defined as the conditional probability of obtaining a t-statistic as large or larger than what was obtained being less than 5 percent (or whatever value we set alpha at) if the null hypothesis is true.

13 For those with a background in statistics, you may recall that we derive the p-value from the t-statistic that is then compared to the t-distribution. We will ignore the t-statistic in this manual because it is the p-value that is used to declare if a difference is statistically significant.

Note: When you read your Excel spreadsheet you will sometimes get a number that looks like:

6.35986E-09

At first glance you might think that this p-value is greater than 0.05 because 6.35 is much greater than 0.05. However, notice that at the end of that number is the notation **E-09**. This indicates that there is a negative exponent. When you have a negative exponent you need to **move the decimal place** as many places to the **left** as the number following the E. So:

6.35986E-09 is really: 0.000000006355986

In this example the p-value is less than 0.05 and you have a statistically significant difference.

Usually we round p-values to just two decimal places. There is no need to get excited about very small p-values. The issue is whether or not it is less than 0.05—how far less than 0.05 does not provide meaningful information. The only concern is whether it is less than 0.05 and, therefore, is significant.

CASE EXAMPLE

CRCC Paired Samples T-Test

As shown earlier, CRCC staff noticed that the scores did increase each time.

Pre-survey average = 3.2
Post-survey average = 4.0
Follow-up survey average = 4.1

Remember: For all of these questions higher numbers indicate greater endorsement of gender equity.

However, they needed to make sure whether these differences were "real" (i.e., statistically significant). They were especially suspicious about the change from post-survey to follow-up survey because that difference was so small.

So CRCC staff looked at the p-values they calculated in their paired samples t-tests (see Page 70 for the Excel results). They found:

Pre- to Post-survey: p less than 0.05
Post- to Follow-up survey: p greater than 0.05
Pre- to Follow-up survey: p less than 0.05

Interpretations: Combining the averages and t-test results, CRCC staff concluded that:

- Students significantly increased in their endorsement of gender equity from the pre-survey to the post-survey. This is what they hoped for and was seen as a positive result.
- There was no significant change from the post-survey to the follow-up survey. This meant that students maintained their beliefs about gender equity. While CRCC staff would have liked to see further increases in gender equity, they were pleased that as time went on students did not revert back to their old ways of thinking.

TASK 4: AVERAGES: DIFFERENCES BETWEEN GROUPS

Keeping in mind the explanation of averages (and the cautions about interpreting them) on Page 109, we can also interpret our results around differences between groups. Specifically, in evaluating prevention programs you may be interested in:

- Were all groups the same at the start of the program?
- Did some groups change more than others?

To answer these questions, start by looking at the averages we calculated in our data analysis phase. This time, however, we are looking at the group averages we calculated as a preliminary step to our independent samples t-tests.

CASE EXAMPLE

CRCC Average Gender Differences

CRCC staff thought it was possible that girls and boys might differ in their endorsement of gender equity in dating and heterosexual relationships. Therefore, they started by looking at the gender-group averages for each of the surveys.

The averages were as follows:

	PRE-SURVEY	POST-SURVEY	FOLLOW-UP SURVEY
BOYS	3.0	3.7	3.9
GIRLS	3.3	4.1	4.2

Remember: For all of these questions higher numbers indicate greater endorsement of gender equity.

Interpretations: What CRCC staff saw:

- On <u>all</u> surveys the boys, on average, scored lower than the girls.
- The gender difference was <u>greatest</u> on the post-survey.

TASK 5: INDEPENDENT SAMPLES T-TESTS

We cannot assume that the differences we see in the averages represent "real" differences. These may be minor differences due to random fluctuations. Therefore, we need to look at the p-values we found in the **independent samples t-tests**.

This is the number that appeared in your worksheet after you ran the independent samples t-test. You have a separate p-value for each comparison you made. In the CRCC example there were three p-values calculated: (1) gender differences on the pre-survey, (2) gender differences on the post-survey and (3) gender differences on the follow-up survey.

- If the p-value is **less than 0.05**, we say that the difference is **statistically significant**. In other words, there really was a difference.
- After determining that there is a significant difference we can look at the **averages** to see which group scored higher. Do not assume which group scored which way! Always go back and check the averages.
- If the p-value is **0.05 or greater**, we say that the difference is **not** statistically significant. In other words, there was no real difference between the groups. Any difference between the means.

Review Page 85 for how to read negative exponents.

CASE EXAMPLE

CRCC Independent Samples T-Tests for Gender Differences

As shown earlier, CRCC staff noticed that boys consistently scored lower than girls:

	PRE-SURVEY	POST-SURVEY	FOLLOW-UP SURVEY
BOYS	3.0	3.7	3.9
GIRLS	3.3	4.1	4.2

Remember: For all of these questions higher numbers indicate greater endorsement of gender equity.

However, they need to make sure whether these differences were "real" (i.e., statistically significant).

So CRCC staff looked at the p-values they calculated in their independent samples t-tests (see Page 82 for the Excel results). They found:

Gender difference on pre-survey:	**p greater than 0.05**
Gender differences on post-survey:	**p less than 0.05**
Gender differences on follow-up survey:	**p less than 0.05**

Interpretations: Combining these two sets of results, CRCC staff concluded that:

- At the start of the program (pre-survey), girls and boys were similar in their attitudes about gender equity in dating and heterosexual relationships.
- On the post-survey and follow-up survey, boys reported significantly less gender equity than girls.

TASK 6: DECIDE IF FOLLOW-UP ANALYSES ARE NEEDED

When you look back over all of your results you may find that they **raise more questions**. Brainstorm any new questions and see if they can be answered by further analysis of the data. The follow-up analyses you run (if any) will depend on the actual results, what you are interested in and if you have the data available to answer your new questions.

CASE EXAMPLE

CRCC Follow-up Analyses

Because the boys were significantly lower than the girls at both the post-survey and follow-up survey, CRCC staff wondered whether the program was actually effective for boys. They noted that the averages for boys did also show change over time, but they wanted to make sure that those changes were significant. Maybe the earlier paired samples t-tests that showed positive change over time were overly influenced by the girls whose data overshadowed the boys.

To make sure that boys showed significant improvements, they ran paired samples t-tests only on the boys. These results showed that for boys:

Pre- to Post-survey:	p less than 0.05
Post- to Follow-up survey:	p less than 0.05

To be thorough, they also ran paired samples t-tests only on the girls. These results showed that for girls:

Pre- to Post-survey:	p less than 0.05
Post- to Follow-up survey:	p greater than 0.05

Interpretations: Combining these results with the gender-specific averages (see previous page), CRCC staff refined their overall conclusions:

- Boys increased their endorsement of gender equity between the pre-survey and the post-survey. They showed further increases between the post-survey and the follow-up survey.
- Girls also increased their endorsement of gender equity between the pre-survey and the post-survey. However, after that they simply maintained their attitudes. For the girls there were no significant increases or decreases between the post-survey and the follow-up survey.

TASK 7: SUMMARIZE AND REPORT YOUR FINDINGS

Congratulations! If you followed these steps you now have a set of results you can use and share with others. To make them useful to your agency and a meaningful contribution to the field, it is important to summarize and report your findings.

A few guidelines can help you complete this final task[14]:

- Write a **concise summary**, in plain language, for yourself. Try to keep your summary to one page. Avoid using too many numbers in this summary. Focus on the main ideas. Make note of both what your findings were and the unanswered questions you have.
- After that, decide who your **audience** for your report will be and what your goals are for sharing the results with them. Even if you are only using your evaluation internally, what you share with frontline staff may be different from what you share with your executive director or board of directors.
- A thorough **evaluation report** will usually include:
  - background information on your program (e.g., program goals and objectives, target population, setting, curriculum, intensity, etc.)
  - description of the evaluation (e.g., evaluation design, measures used, data collection procedures, etc.)
  - key evaluation findings (e.g., changes over time, differences between groups; here you will use the numbers from your results)
  - conclusions, suggestions and recommendations based on the findings
- Do **not** include every detail of your results in the report. Rather, highlight the findings that are most relevant to your audience.
- When possible use graphs, charts, tables or diagrams to illustrate your findings.
- Do not hide **negative findings**. It is very valuable to learn what parts of your program are not working. It is only with this information that we can improve our programs. When reporting negative findings, discuss why you think that was the case and what you will do different in the future to try and achieve your goals.
- Do not downplay your **positive findings**. Celebrate your successes. Be outspoken about what you achieved.

---

14 Adapted from *Introduction to Evaluation Training and Practice for Sexual Assault Prevention* (1999). Published by the Michigan Public Health Institute. Available from Rebecca Campbell, rmc@msu.edu.

- Write your report in a **logical** sequence. Make it easy for a reader who knows nothing about your program or your evaluation to follow.
- **Format** your report (whether it is a written document or slides you use in a presentation) in a way that makes it easy to read and understand. Use italics, bold, underlining and spacing to highlight key points and findings. When applicable, use bullets or numbers instead of long paragraphs. Use headings and sub-headings to help your audience easily follow your train of thought.
- **Edit** and proofread your report. Put yourself in your audience's shoes. Delete any information that is not essential for them to know. Use language that will be easily understood. Make sure the report flows. Use the spell-check and grammar-check tools available on your computer.
- **Share** your report in a timely manner.

CASE EXAMPLE

CRCC Summary of Findings

The following is the summary CRCC staff wrote for themselves to help them think about their findings. This summary was later used to prepare their presentations to their staff and board of directors.

We evaluated our teen primary prevention program by giving students a survey at the start of the summer camp (pre-survey), at the end of camp (post-survey) and again at the end of the school year (follow-up survey). Surveys were given in-person during our regular group meeting times. The evaluation was based on all students who completed all three surveys (20 students). There were three students who dropped out of the program early in the school year. Their data was not included in the evaluation.

Although the program has many goals, this year we evaluated only one of our goals: changing students' attitudes about gender roles. We hoped that students' attitudes would increase in the direction of supporting gender equity in dating and heterosexual relationships.

The evaluation results indicate that we were successful in achieving this goal. At the start of summer camp students showed a wide range of attitudes on the survey. Some students' answers reflected a lot of reliance on gender stereotypes whereas other students already supported gender equity in at least some areas. Students' attitudes depended somewhat on the particular behavior we asked them about.

Our main question was whether students did, in fact, come to endorse gender equity more over the course of the program. We found that they did. There was a significant increase in the average score on the gender equity scale from the pre-survey to the post-survey. This indicates that they did change their attitudes over the course of the summer camp. There was no significant change from the end of camp to the end of the school year. However, we see this as a positive finding because it means that students maintained their positive attitudes over the course of the year.

Unanswered question: Did the monthly meetings during the school year help them maintain their attitudes? Or would they have maintained them even without the meetings? To answer this question we would need to compare two versions of the program: (1) summer camp only and (2) summer camp plus monthly meetings. We might be able to do this in the future if we expand the summer camp to include more students from different schools.

We also wanted to know if girls and boys both benefitted from the program so we compared their average scores on each of the surveys. We found that at the start of the program boys' and girls' attitudes were <u>not</u> significantly different. This means they were equivalent to start. However, on the post- and follow-up surveys boys scored significantly <u>lower</u> than the girls.

(continued on next page)

This finding made us wonder whether our earlier findings about the program having positive effects over time was, in fact, true for both boys and girls. Or maybe it was only the girls who changed, and they changed so much that they were the driving force behind the positive results we saw over time. So we went back and ran the tests again for changes over time. During this second run we looked at boys and girls separately. We found that both boys and girls did, in fact, change in positive directions over time. The patterns were different, however. Girls' attitudes significantly improved over the course of the summer camp but there were no further improvements over the course of the school year. However, boys' attitudes significantly improved over the course of the summer camp and then continued to improve over the course of the school year.

We conclude that the program did significantly increase support of gender equity in dating and heterosexual relationships for both girls and boys. Therefore, we have evidence to support that our program is effective in achieving this goal.

Unanswered question: Could it be that teenagers during this age would change their attitudes anyway simply as part of their normal development? To be sure that it was our program and not something else that caused these changes we would need data from a comparison group of similar teenagers who do not participate in our program. Then we could compare the teens in our program with their peers and see if our teens' attitudes are more positive. This would make our evaluation stronger, but we don't think our agency has the resources to do this type of evaluation.

Unanswered question: Does changing attitudes actually change the students' behaviors in their relationships? Does it change whether or not they speak out against gender inequity when they see it? We cannot assume that changing attitudes changes behavior. So next year we plan on expanding our evaluation to include questions about how the students behave in their relationships and how often they have reacted when they witness gender inequity in relationships.

CHECKLIST

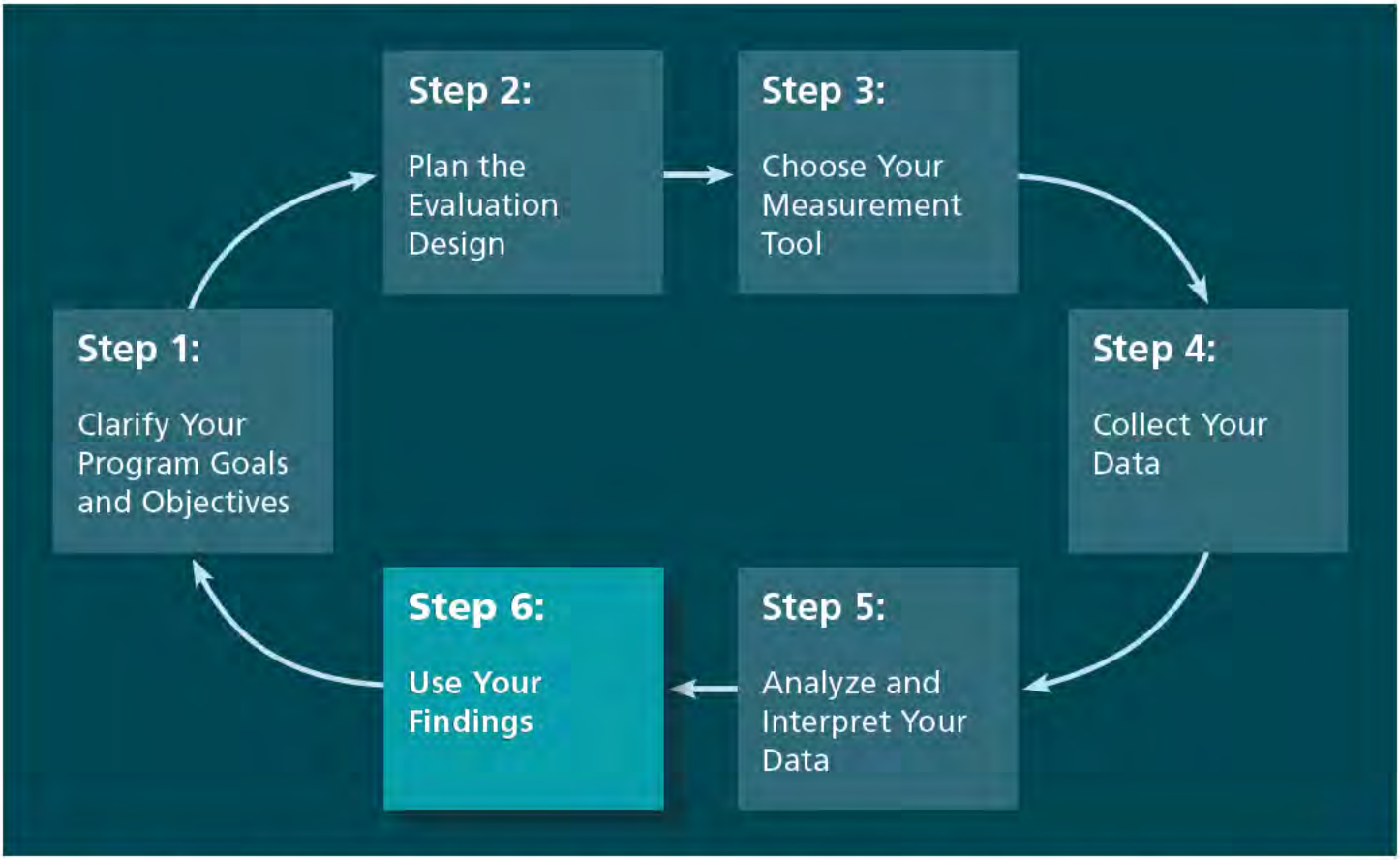
In this section you learned how to interpret your results.

- Interpret Frequencies and Percentages
- Interpret Averages for Changes Over Time
- Interpret Paired Samples T-Tests for Changes Over Time
- Interpret Averages for Differences Between Groups
- Interpret Independent Samples T-Tests for Differences Between Groups
- Conduct Any Follow-up Analyses Necessary and Draw Final Conclusions
- Summarize and Report Findings

USING QUANTITATIVE EVALUATION DATA

- Where Do We Go From Here?
- What?
- So What?
- Now What?

WHERE DO WE GO FROM HERE?



Getting to the results is not the final goal of an evaluation. Rather, it is the first step in strengthening and sustaining your prevention work. Now that you have the results of your evaluation, you will need to think about how to use them. One way to frame the process of using your results is to ask the questions, "What?" "So what?" and "Now what?"[15]

15 This section is adapted from Campbell, R., Greeson, M., Karim, N., Shaw, J., & Townsend, S. (2011). *Evaluating the work of Sexual Assault Nurse Examiner programs in the criminal justice system: A toolkit for practitioners.*

WHAT?

"What?" refers to the actual results from the evaluation that you interpreted in the previous section of this volume and that you summarized as the last step in the previous section. Whether what you found was what you expected or hoped for, remember that there are many factors that can affect the impact of your prevention strategies.

Remember the social-ecological model of sexual violence (see Volume 1). The causes of sexual violence occur at many levels and are woven into the fabric of our society. No single prevention strategy can impact every one of those causal factors simultaneously, nor can any set of strategies that you may be using in your community. This is multi-generational work. Additionally, there may be practical and logistical issues you faced that were out of your control.

SO WHAT?

The next step is to ask, "So what?" What do these results mean? Why do they matter? To unpack these questions, it is important to collaborate with others within your program and your community. Gain as many perspectives as you can on what the findings mean. As you share your findings with others, you may discover new ways of looking at them. The following are some suggestions for sharing your findings with others.

Collaborating with Program Staff, Board and Volunteers

It is important that people who work on your prevention programs as well as staff hear about the results of the evaluation and participate in interpreting the findings. Talking about and working through the evaluation findings within your program first allows you to formulate a preliminary interpretation before sharing them with external parties. This is also an excellent opportunity to engage your entire agency in your prevention efforts. For example, counselors and advocates may not have the time to contribute to the day-to-day planning and implementation of your initiatives, but they may have useful and unique insights on why people may have answered evaluation questions in the way they did.

The following are some ideas to consider when sharing and interpreting the evaluation results with your staff, board and volunteers:

- **The session can be informal or formal.** You should determine which fits best for your program and its usual way of doing things.
- **You may want to have multiple sessions.** You may want to have initial sessions with people directly involved with your prevention work first and later sessions that include people who do not directly work in the prevention arena. Who you include and when will depend on the dynamics in your organization. Also, it may not be possible to fully interpret your findings in one session. It may be helpful to have multiple sessions attending to different topic areas or revisit the same topic in a series of sessions.
- **Create a safe space.** You want to create a space where people feel comfortable speaking out, feel they are being heard by the group and feel that they are making a contribution. Establish ground rules that emphasize respect for one another and for the efforts that went into the prevention strategy and evaluation. You do not want your prevention staff to feel like they are on the "hot seat," especially if the findings were disappointing.

- **Prepare questions that will help guide the interpretive process.** The session leader should have a list of questions to guide the discussion. Remember to consider not only if there were positive changes, but also whether the size of that change is what you were hoping for. Some questions to consider are:
  - Were any of the results surprising? What and why?
  - Did the results confirm anything we already suspected? What and why?
  - Why do we think we had the results we did?
  - Are we proud of any of the results? If so, what will it take to maintain them in the future?
  - Are we disappointed in any of the results? If so, what will it take to change them in the future?

Collaborating with Community Partners

After talking about the findings with people in your program, it is time to collaborate with community partners. The purpose of sharing your findings with community partners is two-fold.

First, your partners will provide additional perspectives as to why you found what you did. Second, their collaboration can help build and strengthen your relationships with partners.

In addition to the considerations listed above for collaborating within your program, when sharing evaluation findings outside of your program you may also want to consider the following:

- **Write a concise summary of your program's interpretation of the findings**. Try to keep your summary to a couple of paragraphs. Avoid using too many numbers. Focus on the main ideas. Make note of both what your findings were and unanswered questions.
- **Determine the audience and session goals.** You may decide it is best to present the findings to each stakeholder group separately (e.g., educators, social service agencies, police, medical personnel, etc.) or to present the findings to all stakeholders in the same session. Be sure that the session goals are tailored to the audience's interests, needs and expertise.
- **Provide all relevant information.** Your community partners likely will not be able to make meaningful contributions to the discussion if they do not have all the relevant information. Be sure they understand:
  - Background information on the prevention program (e.g., goals, participants, setting, etc.)
  - Description of the evaluation (e.g., how you did the study, what the survey asked, etc.)
  - Key evaluation findings (e.g., what you learned)
- **Do not hide unexpected findings or downplay the disappointing findings.** One purpose in collaborating with community partners is to identify where the shortfalls are and to identify opportunities for strengthening your work. Hiding unexpected or disappointing findings will prevent this from happening. Additionally, it is important to celebrate success. Be sure to acknowledge what you and your community partners are doing well.
- **Use visuals.** Use graphs, charts, tables or diagrams whenever possible to illustrate your evaluation process and findings.

- **Develop additional questions to guide the interpretive process with community partners.** Some questions to consider include:
  - How does each community partner think they contributed to the findings?
  - How can community partners expand on what you achieved in their own work?
  - What can each community partner do to strengthen the impact of the program in the future?
  - How can community partners use the findings to support their own work?
  - What parts of the evaluation should be shared more widely in the community? With whom and how?
- **Share your findings in a timely manner.** While it will take time to prepare your findings and organize a session with community partners, it is best to do so when the evaluation process and findings are still fresh and relevant.

After interpreting your results along with your staff, board, volunteers and community partners, it is time to take action. In the "Now What?" phase of the process, you will use the interpretations of the findings to take further action.

NOW WHAT?

An evaluation is only valuable if you use the findings. The goal of this final phase is to use your findings and how you interpreted them to inform your ongoing prevention work. There are a number of key areas where you can do this.

Strategic Planning

Use your evaluation findings to make informed decisions about what you will do in the future. Consider what the findings indicate about your strategic plans:

- Does this program help to meet the **goals** of your agency's strategic plan?
- Is this specific prevention strategy worth **continuing**?
- Would it be beneficial to invest **more resources** into it? From where might those resources come?
- Can it be continued with **fewer resources**? If so, where would you reallocate resources?
- Was there enough success that you want to expand this strategy? If so, where would you take it?
- Do the findings indicate any **gaps** in the achievement of your goals that need to be addressed in new ways?

Strengthening Community Relationships

Even when communities support our work, they often do not have a full understanding of what we do and why it is important. Consider if there are ways you can use your evaluation findings to build understanding of and support for your prevention work:

- Do the findings validate the need for **continuing your existing relationships**? Most directly, do they speak to the need for those who collaborated or cooperated with you on the prevention strategy to continue working with you?
- Can your findings be used to **expand existing relationships** to include collaborative work on prevention and not only on services?
- Are there stakeholders in the community who are not currently working with you who might be excited about these findings? Can you use the findings to **engage new partners**?

Shape Further Training

Sometimes evaluation findings can shed light on specific areas where further training and skill-development is needed. Training can result in greater impacts of the prevention program. For example:

- If there were specific areas of the evaluation that showed less impact than others, are those areas where further training for **your staff** is needed? Perhaps they need to develop their understanding of and skills for engaging the participants in those parts of the program.
- Could training for **community partners** increase the impact of the program or close the gaps? This is especially important to consider for school-based prevention programs. If there is an area where impact was lacking, perhaps training teachers and staff on how they can engage students around those issues in their own roles will help to transfer the lessons of the prevention program into students' daily lives.
- If there were successes in the program, those may also indicate areas for training a **broader array of community members**. For example, if you found that the program facilitated by your staff had a lot of impact on participants' likelihood of intervening as empowered bystanders, perhaps that is an area where you want to train more leaders in the community on how to teach and reinforce those skills in their own roles. This will increase the saturation of the community with skills and strategies that you found to be effective.

Seek Support and Funding

Evaluation findings can be used to seek support and funding for your prevention programming. The fact that you engaged in a systematic evaluation of outcomes will make your work far more attractive to potential funders. Increasingly, funders are looking for "evidence-based practices." Your evaluation is evidence, so use it! This is true, regardless of whether the findings showed impact or not.

- If you found that your program had positive impacts on participants' knowledge, attitudes, intents and/or behaviors, use those findings to justify why and how your program is working and to demonstrate that it is **worth investing funds in to continue or expand**.
- If you found areas where your program is <u>not</u> having the impact you wanted, you can also use those findings to demonstrate why you need **support to improve** those aspects of your work. For example, if you found that you increased knowledge about sexual violence but did not achieve changes in participants' likelihood of intervening as empowered bystanders, then you might use that information to ask a school district to give you more time with students so you can spend more time developing those skills, or you might ask a foundation for funds to support more staff training in that area or to fund development of a longer program that may be more effective at the skill-building.

When using your evaluation to write grant applications, the following strategies can increase the chances of receiving funding:

- **Relate what you are doing (or want to do) to the models of sexual violence prevention.** Do not assume the funders know anything about sexual violence prevention. Explain how what you are doing relates to the complex, multi-level causes of sexual violence. Explain why changing knowledge and attitudes does not change behaviors. Demonstrate why it is important to focus on skills and behaviors that are preventive. Distinguish between prevention vs. risk reduction vs. awareness.

- **Relate what you are doing (or want to do) to the existing research literature.** While we lack robust evidence-based practices, we do have an emerging body of literature that supports some practices such as bystander empowerment, social norms campaigns and multi-session programs. Show how your evaluation fits into and expands upon that research. If you need help accessing research literature, call the National Sexual Violence Resource Center (877-739-3895) and they can send you relevant articles and reports.
- **Clearly explain all of the activities you have engaged in to evaluate your program.** Describe how you collected your data and present the findings. If you have done other evaluations, present those, too.
- **Avoid technical jargon.** Whenever possible, use language that people outside the field can understand. If you do use technical terms (e.g., "social-ecological model," "bystander empowerment," "risk reduction," etc.), define them.
- **Be consistent.** If you use a phrase in one place in your application, continue to use that same phrase throughout the application. Inconsistency creates confusion.
- **Include a plan for future evaluation.** Even though you completed an evaluation using this resource kit, you need to continue evaluating your work. Explain how you will evaluate what you do in the future.
- **Proofread!** Proofread your own writing and then have at least two people who did not contribute to the writing proofread it. Any errors, even simple typos, reflect poorly on your program. Build enough time into your writing process that you can set aside your "final" application for a few days. Then take a final look at it before sending it off. You may be amazed at what you see after you step away from it for a while.

Develop Ongoing Evaluation Processes

Finally, use what you learned from this evaluation to shape how your agency approaches evaluation more broadly. Evaluation should be an ongoing process. When it is integrated into your programming and becomes part of the habit or way of doing business for your agency, your staff will continue to develop their skills and it will take fewer resources to do evaluation.

Just like agencies have worked hard to become "survivor-centered," "trauma informed," etc., they can work to become "data informed." This requires:

- Looking at how evaluation is **supported** (funds, staff, volunteers, time, etc.)
- Examining how it is included and used in **strategic planning**
- **Valuing** and **using** the information it provides
- Having an **infrastructure** that makes evaluation easier to do and consistently done
- Considering evaluation **from the beginning** of a project rather than adding it at the end
- Evaluating your agency's work **even when a funder does not require it**

RESOURCES

Primary Prevention

Primary Prevention and Program Evaluation Technical Assistance Guide and Resource Kit. Multi-volume toolkit available from the Pennsylvania Coalition Against Rape.

Sexual Violence Prevention: Beginning the Dialogue. Report available from the Centers for Disease Control and Prevention.

Guidelines for the Primary Prevention of Sexual Violence and Intimate Partner Violence. Report available from the Virginia Domestic and Sexual Violence Action Alliance at www.vsdvalliance.org.

Prevention Connection. Online resources, web conferences, online training modules and listserve available at www.preventconnect.org.

Program Evaluation / Research Methods

Getting to Outcomes for Sexual Violence Prevention. Manual forthcoming from the Centers for Disease Control and Prevention.

Evaluation: A Systematic Approach by Peter Rossi, Howard Freeman and Mark Lipsey. Book available from Sage Publications, 1999.

Community-Based Participatory Research for Health, edited by Meredith Minkler and Nina Wallerstein. Book available from Jossey-Bass, 2003.

Feminist Methods in Social Research by Shulamit Reinharz. Book available from Oxford University Press, 1992.

Fostering Collaborations to Prevent Violence Against Women. Report available from the National Violence Against Women Prevention Research Center at http://www.musc.edu/vawprevention/.

Quantitative Data Analysis

Statistics in Plain English by Timothy Urdan. Available from Lawrence Erlbaum Associates, 2005.

Qualitative Data Analysis

Qualitative Research and Evaluation Methods by Michael Quinn Patton. Book available from Sage Publications, 2002.

Doing Qualitative Research: Circles Within Circles by Margot Ely with Margaret Anzul, Teri Friedman, Diane Garner and Ann McCormack Steinmetz. Book available from Falmer Press, 1991.

Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory by Anselm Strauss and Juliet Corbin. Book available from Sage Publications, 1998.

PENNSYLVANIA COALITION AGAINST RAPE
125 North Enola Drive • Enola, PA 17025
717-728-9740 • 800-692-7445 • TTY 877-585-1091 • pcar.org