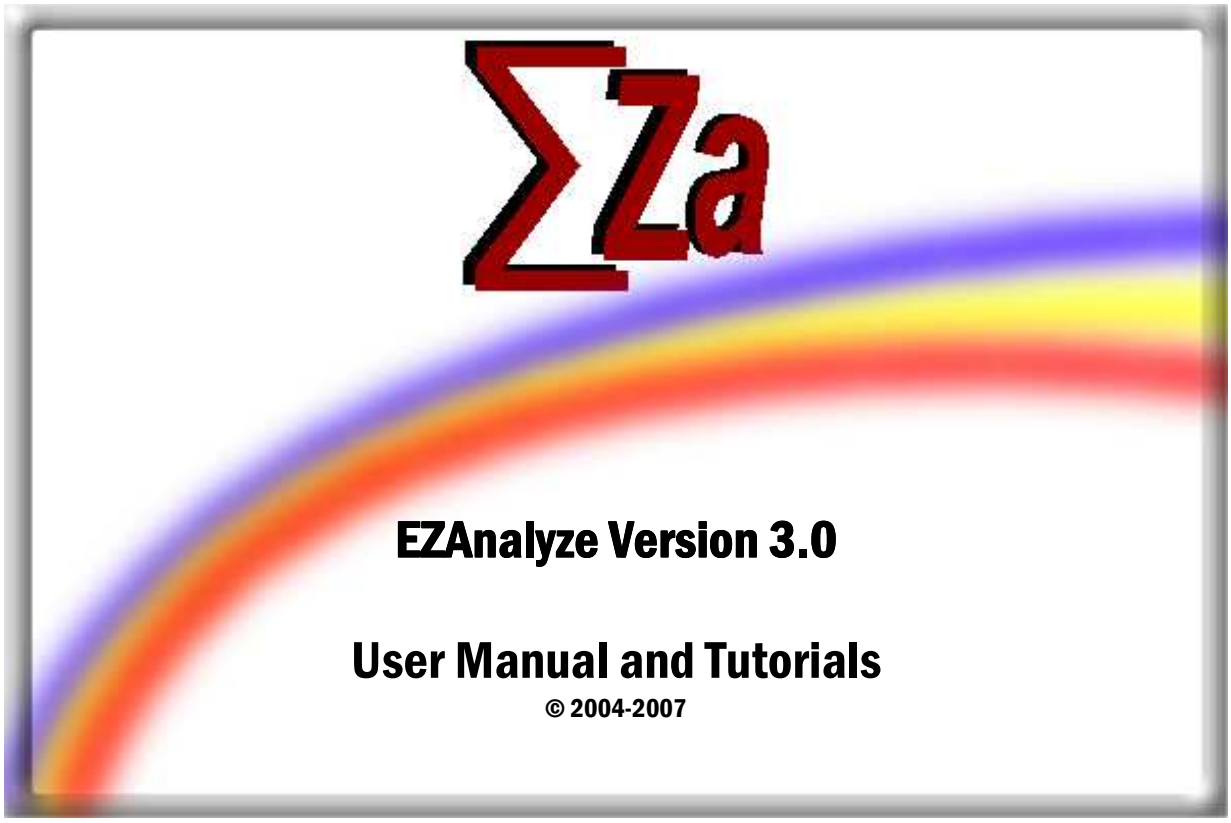


EZAnalyze Version 3.0

User Manual and Tutorials

By Tim Poynton, Ed. D.
Director, School Counseling Program
Suffolk University
Education and Human Services Department
73 Tremont Street, Room 724
Boston, MA 02108

Please Reference Using the Following Citation:

Poynton, T. A. (2007). EZAnalyze (Version 3.0) [Computer software and manual]. Retrieved from http://www.ezanalyze.com

About EZAnalyze

EZAnalyze has been developed to help educators, particularly school counselors, use data efficiently and effectively. The types of procedures included with EZAnalyze are based on the American School Counselor Association's (ASCA) National Model; however, the language and capabilities contained within EZAnalyze are suitable for all educators.

The ASCA National Model calls on counselors to use data in several ways, namely:

DISAGGREGATING DATA. Disaggregating data, simply put, is breaking down data by different categories. For example, the "achievement gap" - the differences observed in academic achievement by ethnicity - is highlighted by disaggregating academic achievement data by ethnicity.

CHANGES OVER TIME. Changes over time can be shown by creating difference variables. For example, if you are interested in knowing how differently students scored on a standardized test in 1999 and 2003, EZAnalyze can create a new score that shows you the amount of improvement or decline for each student. Alternatively, you can use the graphing functions built into EZAnalyze to create a chart depicting how differently students scored in 1999 and 2003.

RESULTS REPORTS WITH SIMPLE STATISTICS. The types of statistics needed to effectively create a results report do not need to be complicated. EZAnalyze makes it easy to compute simple "descriptive statistics" on your data, such as means, standard deviations, percentages, and a few others.

ADVANCED FEATURES. The advanced features of EZAnalyze are correlation, t-tests (one sample, paired, and independent samples), ANOVA (single factor between subjects and Repeated Measures), chi square, and Regression. These features are “advanced” because they are above and beyond what is needed in to accomplish the goals of the ASCA National Model, and require knowledge of how to conduct and interpret statistical hypothesis tests to use them appropriately. HOWEVER, in keeping with the theme of the software, these advanced features are just as easy to use as the other features.

Index

Using the EZAnalyze Manual and Tutorials

Most of the information available in the "Manual" section of this document is also available in the EZAnalyze Help files. Once you have successfully installed EZAnalyze, you are encouraged to follow the tutorials to develop an understanding of how to prepare yourself and your data for EZAnalyze.

Installation Instructions

Installation and Upgrading Instructions - Version 3.0 Automated Installer

Once you have saved the file (EZAnalyze3.xls) to your computer from the website:

1. Close any Excel files you may have open
2. Open the EZAnalyze3.xls file
3. Indicate that you trust the source, and click 'Enable macros' or 'Enable this content'
   (Excel 2007 only - click on the 'Options' button that appears under the ribbon in the 'Security Warning' to display the dialog box, and select 'Trust all documents from this publisher' or 'Enable this content')
4. Follow the on screen instructions

NOTE: This file is digitally signed by me under the name EZAnalyze. When you open the file, there is a long pause while the digital signature is verified. This helps protect you.

If you are upgrading from a previous version of EZAnalyze, this will automatically uninstall the older version

If you do not see the 'Enable macros' dialog box, instructions should be visible in the EZAnalyze3.xls file to change your security settings. It might be easier to download EZAnalyze3.xla (no automated installer) and follow the 'manual' installation instructions below...

Installation and Upgrading Instructions - Version 2.5 and 3.0 without Automated Installer

Save the file (EZAnalyze.xla or EZAnalyze3.xla) in a directory on your computer (preferably NOT your desktop – create a new folder in your Documents folder).

Once you have saved the file to your computer from the website:

1. Open Excel and go to the "Tools" menu
   (Excel 2007 only - go to the 'Office Button' in the upper left corner and click the "Excel Options' button)
2. Select "Add-Ins" from the Tools menu
   (Excel 2007 only - select 'Add-Ins' from the list on the left, then click 'Go' next to 'Manage Add-Ins')
3. Click on the "Select" (Mac) or “Browse” (PC) button and go to the folder you saved the "ezanalyze.xla" file to and open the folder
4. Highlight the "ezanalyze.xla" file and click Open
5. Select OK in the Add-Ins dialog box, then click OK again
6. You should now see a new menu called "EZAnalyze©" next to the Excel help menu.

Upgrading Instructions

For manually upgrading from a previous version of EZAnalyze

Once you have saved the program file (EZAnalyze3.xla) to your computer from the website:

1. Open Excel and go to the "Tools" menu
   (Excel 2007 only - go to the 'Office Button' in the upper left corner and click the "Excel Options' button)
2. Select "Add-Ins" from the Tools menu
   (Excel 2007 only - select 'Add-Ins' from the list on the left, then click 'Go' next to 'Manage Add-Ins')
3. Uncheck the box next to EZAnalyze and click OK. You should see the EZAnalyze menu disappear.
4. Close and restart Excel
5. Repeat steps 1 and 2.
6. Click on the "Select" (Mac) or “Browse” (PC) button and go to the folder you saved the "EZAnalyze3.xla" file to and open the folder
7. Highlight the "EZAnalyze3.xla" file and click Open
8. Select OK in the Add-Ins dialog box, then click OK again
9. If you get a message asking if you would like to replace the existing file at "...\ezanalyze.xla", click Yes.
10. You should now see a new menu called "EZAnalyze©" next to the Excel help menu.

Overview of EZAnalyze

EZAnalyze is an "add in" to Microsoft's Excel. Microsoft Excel is a powerful tool for organizing, displaying, and analyzing data. However, the formulas to perform the data analyses and make sense of your data are not that easy to use in Excel. EZAnalyze "bridges the gap" between you, the user, and Excel's powerful data analysis features by adding a new menu to Excel - the "EZAnalyze" menu. From the EZAnalyze menu, you can point and click your way to making sense of your data quickly and easily.

The way EZAnalyze generally works is that you:

1. Open a data sheet in Excel and make sure that it is compatible with EZAnalyze. See “Setting up your data sheet” in the Manual, or participate in “Tutorial 1” to learn more.

2. Select a function from the EZAnalyze menu (available functions are Describe, Disaggregate, Graph, New variable, or one of the Advanced functions – see associated topics in the Manual to learn more, or participate in Tutorial 2.

3. Provide EZAnalyze with the names of the variables you would like the function to work on.

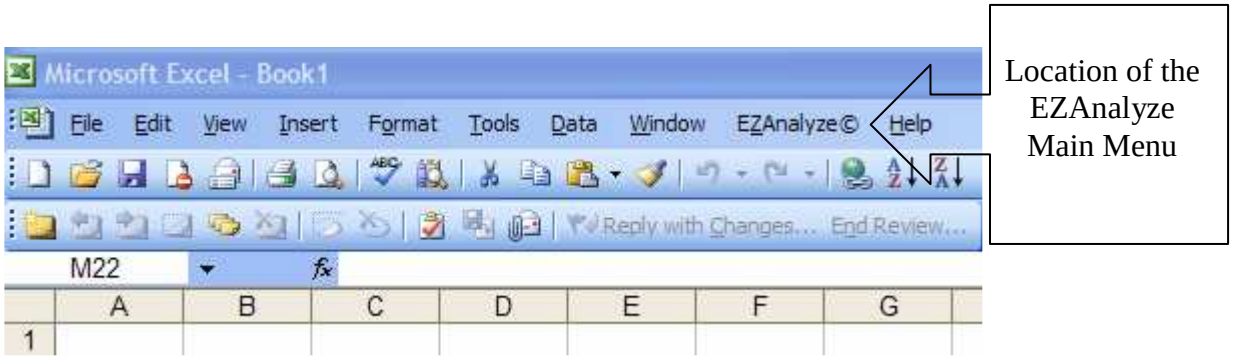
4. Look at the Results Report to see the results of your data analysis. Learn how to interpret the various EZAnalyze results reports by reading the Manual section on “Interpreting your results reports.”

Each time you select a new function from the EZAnalyze main menu, a new results report is created as a new Excel sheet. To perform other analyses using EZAnalyze, return to your data sheet (usually sheet 1). All of the results from previous analyses will be kept. If you want to save all of your results reports, select "File - Save" from the Excel menu, and all of your results reports and your original data sheet will be saved in the same file for future reference.

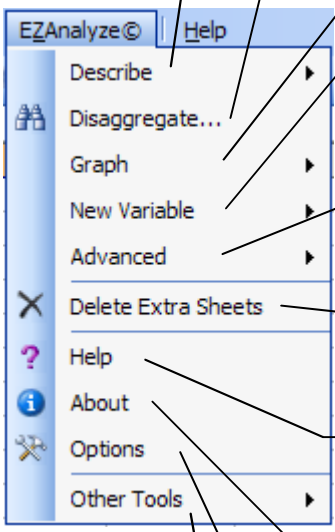
To uninstall EZAnalyze: Go to the Options menu of EZAnalyze and click on the Uninstall button. You can now delete the EZAnalyze files if you desire.

The EZAnalyze Menu

To locate EZAnalyze, open Excel; you will see EZAnalyze next to the help menu.



The EZAnalyze Menu – for more details about each function, participate in the tutorials or view the related topic in the Manual.



DESCRIBE – Contains the *Percentages* and *Descriptive Statistics* functions

DISAGGREGATE – Contains the *Disaggregate* function; permits sorting of results into categories

GRAPH – Contains the *Histogram, Disaggregate* and *Multiple Variable* graphing functions

NEW VARIABLE – Contains the functions to create new *Summary*, *Difference*, *Standardized, Percent Change, Percentile Rank, Random Number, and Binary* Variables

ADVANCED – Contains *Correlate, T-Test, ANOVA (Within and Between Subjects), Regression, and Chi Square*

Delete Xtra Sheets – A function to quickly delete all of the sheets EZAnalyze created as you are exploring your data

HELP – Provides access to the extensive help documentation online; the help topics are also available in the Manual

ABOUT – Contains the EZAnalyze version and contact information, License agreement, and the option to Donate

OPTIONS – Provides you with options to change how the menu is displayed, and uninstall EZAnalyze

OTHER TOOLS – Contains the *Amalgamate Data* and *Transpose Data* functions

EZAnalyze Tutorials

Tutorial 1. In this tutorial, you will help "Pat," a high school counselor, set up data for analysis with EZAnalyze. Pat has already distributed and collected a pretest and a posttest to evaluate the effectiveness of a set of classroom guidance lessons, and has the student's grades, but is unsure of what to do next. Pat would like to answer the following questions:

1. How much did student scores increase from pretest to posttest?
2. How much did student grades increase from quarter two to quarter three?

Help Pat put the data into Excel and make some sense of it! By participating in this tutorial, you will learn how to create a data set "from scratch," how to easily create a "summary score" and a "difference score," and how to interpret the EZAnalyze Results Report for Descriptive Statistics.

Specific content

Tutorial 2. In this tutorial, Pat has assembled a much larger data set. All of the data is already in Excel and ready to go – but your help is needed to help Pat analyze and interpret the data. Pat would like to answer the following questions with the data:

1. What are the characteristics of people who responded to the survey?
2. What is the "achievement gap" (the academic performance of each ethnic group) in this high school?
3. How connected do students feel to various school staff members?
4. What do students think about their classes?
5. What is the relationship between feeling connected to school staff and academic achievement?
6. How many students from each ethnic group are there in this sample?
7. Do male and female students have similar academic achievement?

In this tutorial, you will learn **when**, **why**, and **how** to use a few other EZAnalyze functions not included in Tutorial 1, and learn how to read EZAnalyze results reports.

Specific Contents

EZAnalyze Tutorial 1

Pat's Data

Pat conducted a set of 8 classroom guidance lessons, one lesson per week, on study skills. The goals of this intervention were to increase the students' knowledge of study skills and increase the students' academic achievement. To evaluate the study skills intervention, Pat developed a 5-question quiz to measure study skills knowledge, and will look at changes in their grades from the quarter before and after the intervention to measure academic achievement.

Below is the data Pat has for seven students. For each student, Pat recorded if they answered a question correctly or incorrectly on the quiz before the intervention (pretest) and after the intervention (posttest), and wrote their grades for quarter 2 and quarter 3 on the sheets.

Pretest		Posttest		q2 grade	q3 grade
1. wrong	4. wrong	1. right	4. right		
2. right	5. wrong	2. right	5. right	81	88
3. wrong		3. wrong			
Pretest		**Posttest**		**q2 grade**	**q3 grade**
1. wrong	4. right	1. right	4. right		
2. wrong	5. right	2. right	5. right	79	87
3. wrong		3. right			
Pretest		**Posttest**		**q2 grade**	**q3 grade**
1. wrong	4. wrong	1. wrong	4. wrong		
2. right	5. wrong	2. right	5. wrong	87	86
3. wrong		3. right			
Pretest		**Posttest**		**q2 grade**	**q3 grade**
1. wrong	4. wrong	1. right	4. right		
2. wrong	5. wrong	2. right	5. right	62	84
3. wrong		3. right			
Pretest		**Posttest**		**q2 grade**	**q3 grade**
1. wrong	4. right	1. right	4. right		
2. right	5. right	2. right	5. right	80	91
3. wrong		3. right			
Pretest		**Posttest**		**q2 grade**	**q3 grade**
1. wrong	4. wrong	1. right	4. right		
2. right	5. wrong	2. right	5. right	81	84
3. right		3. wrong			
Pretest		**Posttest**		**q2 grade**	**q3 grade**
1. wrong	4. wrong	1. right	4. right		
2. wrong	5. wrong	2. right	5. wrong	54	78
3. wrong		3. right			

Helping Pat see if the study skills intervention was effective!

EZAnalyze Tutorial 1

CREATING VARIABLE NAMES

Above is Pat's data for seven students. Now that all of the data are ready, it is time to put it into Excel to prepare for EZAnalyze to do its part. If you are unsure how to start Excel, you can usually find it by going to the **START** menu, then **PROGRAMS**, then **MICROSOFT OFFICE**. Once you have found the Microsoft Office folder, you should see an icon for Microsoft Excel – click on it to open.

The next task is to put the information from the students into Excel. Before you start just plugging the numbers in, lets first find out what EZAnalyze expects the data to look like. EZAnalyze wants the data to:

A. Have the first row contain the "variables names."
B. Have each row after the first row contain data from one student.
C. Have data that are numbers.

SO, Pat's "data set" needs to first have the variable names identified. By looking at the data, we see that there are a total of 12 different items that will become variables in the Excel data sheet – 5 variables for the pretest questions, 5 variables for the posttest questions, and 2 variables for the grades. *To keep your data organized, it is a good idea to always have an ID number for each student – that way, if a mistake is made while entering the numbers, you can easily find out where the error is.* The total number of "variables" in our data set, then, is 13; 1 for the ID number, and 12 for the student data.

NOW, go to your Excel data sheet and create the 13 variables by giving each variable a name. An example of a data sheet with variable names is below.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	ID	pre1	pre2	pre3	pre4	pre5	post1	post2	post3	post4	post5	Q2	Q3
2													

Your Excel data sheet should look something like this now – row one has the 13 "variable names" in column A through column M.

The first step to setting up Pat's data is done! One more step…

Pretest		Posttest		q2 grade	q3 grade	
1. wrong	4. wrong	1. right	4. right			
2. right	5. wrong	2. right	5. right	81	88	1
3. wrong		3. wrong				
Pretest		**Posttest**		**q2 grade**	**q3 grade**	
1. wrong	4. right	1. right	4. right			
2. wrong	5. right	2. right	5. right	79	87	2
3. wrong		3. right				
Pretest		**Posttest**		**q2 grade**	**q3 grade**	
1. wrong	4. wrong	1. wrong	4. wrong			
2. right	5. wrong	2. right	5. wrong	87	86	3
3. wrong		3. right				
Pretest		**Posttest**		**q2 grade**	**q3 grade**	
1. wrong	4. wrong	1. right	4. right			
2. wrong	5. wrong	2. right	5. right	62	84	4
3. wrong		3. right				
Pretest		**Posttest**		**q2 grade**	**q3 grade**	
1. wrong	4. right	1. right	4. right			
2. right	5. right	2. right	5. right	80	91	5
3. wrong		3. right				
Pretest		**Posttest**		**q2 grade**	**q3 grade**	
1. wrong	4. wrong	1. right	4. right			
2. right	5. wrong	2. right	5. right	81	84	6
3. right		3. wrong				
Pretest		**Posttest**		**q2 grade**	**q3 grade**	
1. wrong	4. wrong	1. right	4. right			
2. wrong	5. wrong	2. right	5. wrong	54	78	7
3. wrong		3. right				

Notice how Pat's Data now has the student ID numbers – they are circled, next to the "q3 grade."

EZAnalyze Tutorial 1

ENTERING DATA

Now that Pat's data sheet has variable names, it is ready for the data. Notice how the variable names are in roughly the same order as they are on the paper? This will make it easier to move the data from the paper to the Excel data sheet.

As you can see from looking at Pat's data, not everything is in a number form – the pretest and posttest questions are listed as "right" or "wrong." We need to come up with *numbers* to represent these words. Let's use "0" to indicate a "wrong" answer, and "1" to indicate a right answer. Using 0 and 1 as the numbers has some distinct advantages over using other numbers – by using 0 and 1, we can easily have EZAnalyze figure out the "total score" for each student by adding up all the questions. While you could do this in your head, why not just let EZAnalyze do it for you? You will learn how to do this later.

NOW we have everything we need to start putting the numbers from the paper into the Excel data sheet. Let's do it!

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	ID	pre1	pre2	pre3	pre4	pre5	post1	post2	post3	post4	post5	Q2	Q3
2	1	0	1	0	0	0	1	1	0	1	1	81	88
3	2	0	0	0	1	1	1	1	1	1	1	79	87
4	3	0	1	0	0	0	0	1	1	0	0	87	86
5	4	0	0	0	0	0	1	1	1	1	1	62	84
6	5	0	1	0	1	1	1	1	1	1	1	80	91
7	6	0	1	1	0	0	1	1	0	1	1	81	84
8	7	0	0	0	0	0	1	1	1	1	0	54	78
9													

Your Excel data sheet should look something like this – one row for the variable names (row 1), one row of data for each student (rows 2-8), and the student data for the 13 variables in the columns (columns A-M).

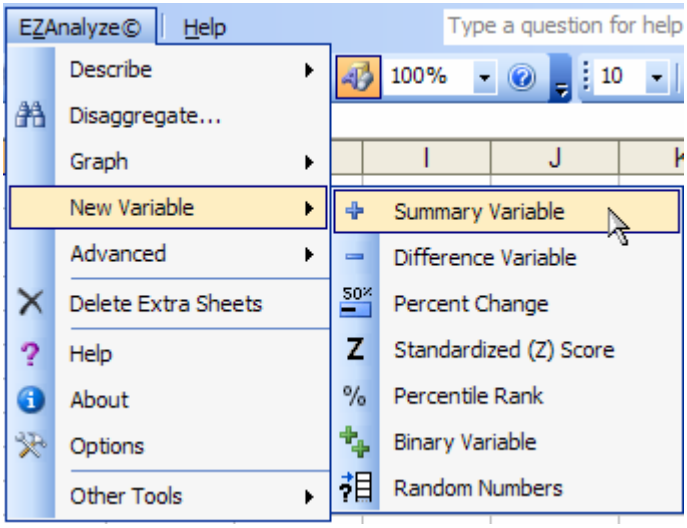
The second step to setting up Pat's data is complete!
Now we can help Pat answer the questions!

CREATING A SUMMARY SCORE

NOW that Pat's data is in an Excel data sheet, we can start putting EZAnalyze to work. The first thing we will want to do is create some new variables to help us make sense of the data.

Create a summary score for the Pretest quiz – we are not really all that interested in each individual question on Pat's quiz. What we are interested in is the total number of questions each student got right – that will give us a total score, which is easier to work with. To create a *total pretest score* variable, follow these steps:

STEP 1



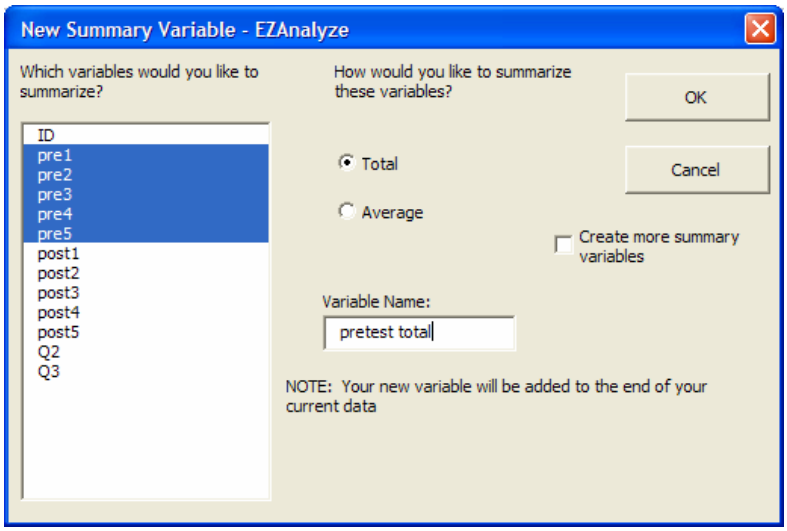
Go to the EZAnalyze Menu and select "New Variable", then select "Summary Variable" from the sub-menu

NEXT

EZAnalyze Tutorial 1

CREATING A SUMMARY SCORE (Continued)

STEP 2



Highlight the pretest variables (pre1 – pre5) by clicking on each one. Make sure "Total" is selected, and give your new variable a name in the "Variable Name" box; click OK.

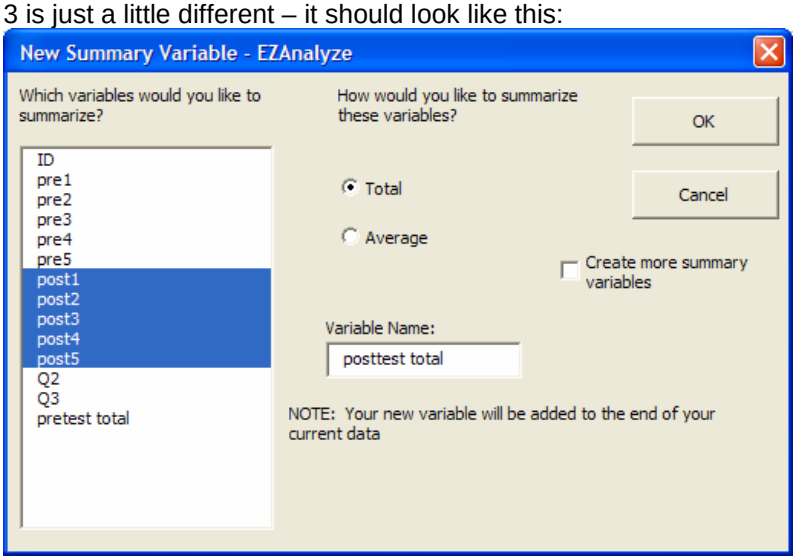
You have created a new variable that gives you the total score – the number of questions right – for the pretest!

You should now see a new variable in column N named "pretest total."

L	M	N
Q2	Q3	pretest total
81	88	1.000
79	87	2.000
87	86	1.000
62	84	0.000
80	91	3.000
81	84	2.000
54	78	0.000

What your data sheet should look like

NOW, follow the same procedure for creating a total score for the posttest – steps 1 and 2 remain the same, step 3 is just a little different – it should look like this:



Creating the posttest total variable…

L	M	N	O
Q2	Q3	pretest total	posttest total
81	88	1.000	4.000
79	87	2.000	5.000
87	86	1.000	2.000
62	84	0.000	5.000
80	91	3.000	5.000
81	84	2.000	4.000
54	78	0.000	4.000

What your data sheet should look like

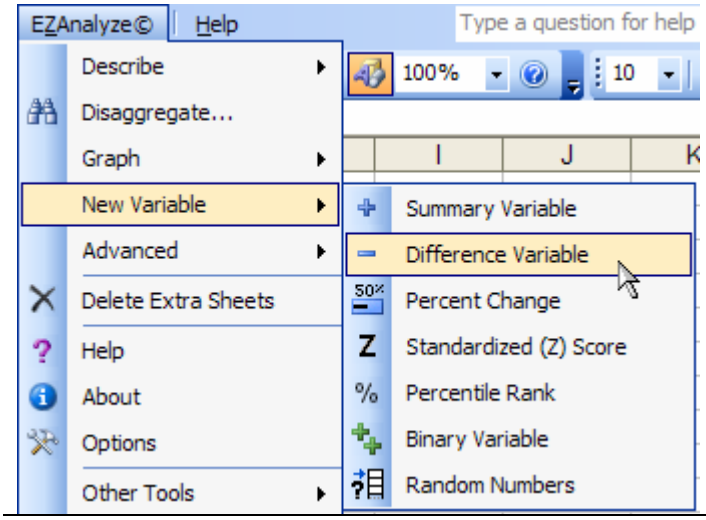
EZAnalyze Tutorial 1

CREATING A DIFFERENCE SCORE

Now that we have created summary scores for the pretest and the posttest, the only other thing that needs to be done is to create difference scores for the pretest – posttest data, and the quarter 2 and quarter 3 grades. By creating a difference score, we will be able to see how much improvement (or decline) has occurred over time for each student.

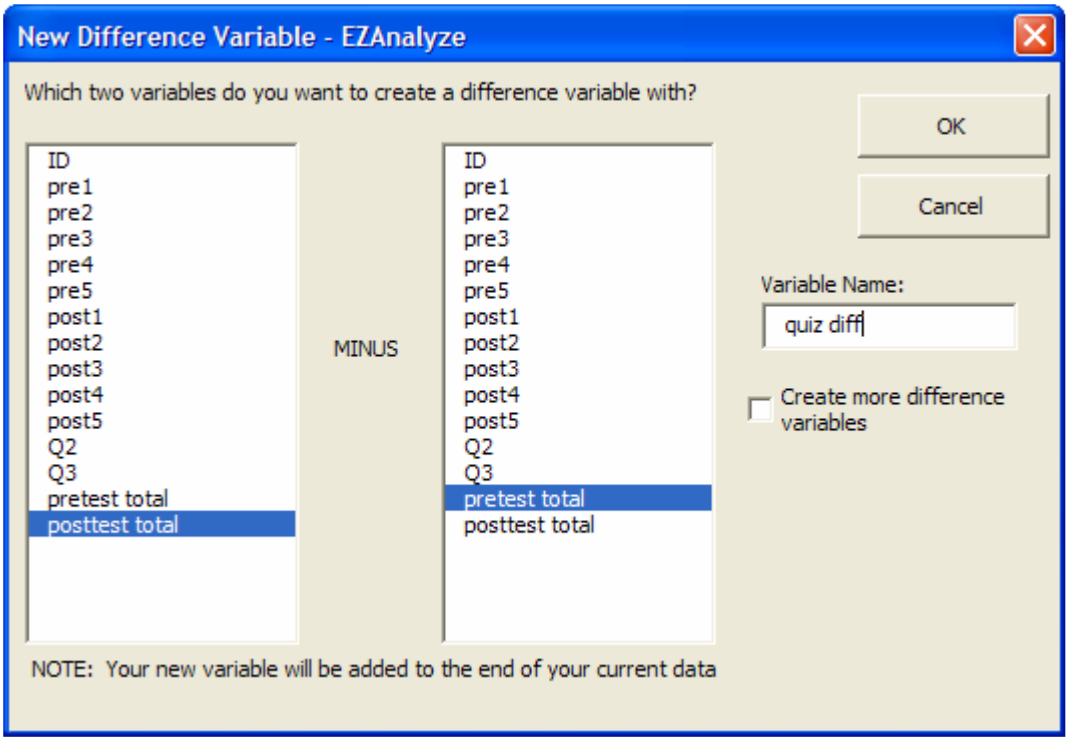
Simply put, a "difference score" is one variable subtracted from another variable for each student. When we subtract the pretest score from the posttest score, the resulting difference score is the amount of improvement or decline from the pretest to the posttest – if the difference score is negative, it indicates that the person's score decreased over time. If the difference score is positive, it indicates that the person's score increased over time. Let's go through the steps to create a difference score, then you will see how EZAnalyze does the math.

STEP 1



Go to the EZAnalyze Menu and select "New Variable", then choose "Difference Variable" from the sub-menu

STEP 2



Highlight the "posttest total" in the first list, , and "pretest total" in the second list. Make sure you give your new variable a name in the "Variable Name" box, and click OK.

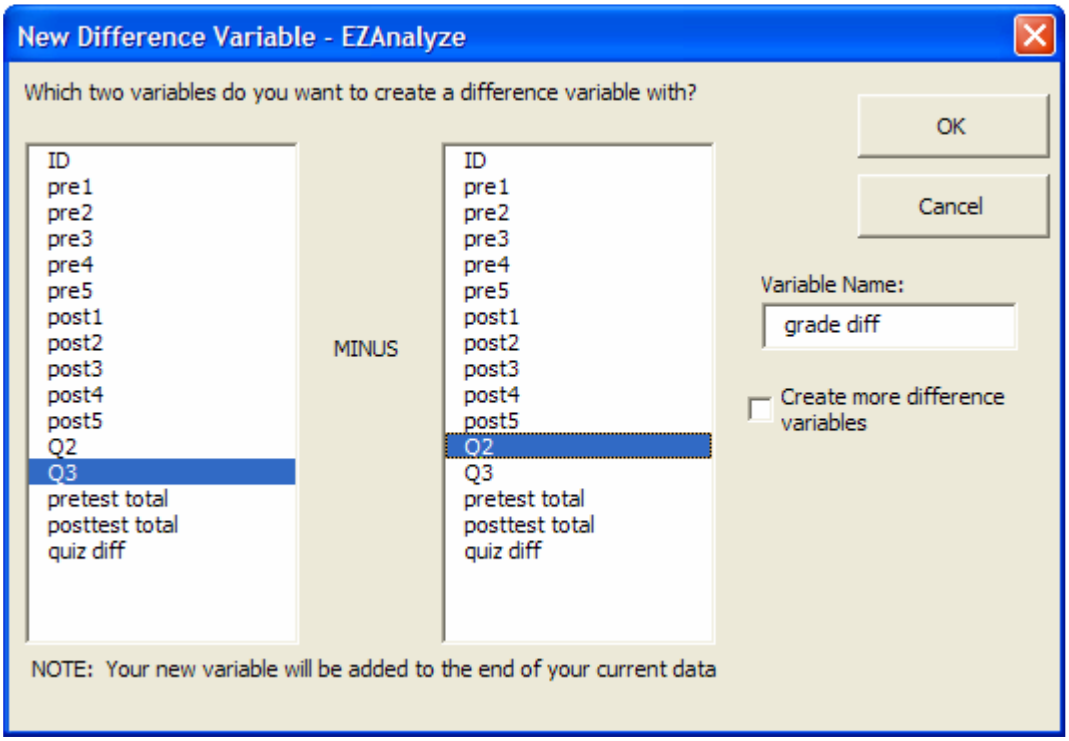
EZAnalyze Tutorial 1

You should now see a new variable in Column P called “quiz diff.” As you can see, the “quiz diff” is equal to “pretest total” subtracted from the “posttest total”

N	O	P
pretest total	posttest total	quiz diff
1.000	4.000	3.000
2.000	5.000	3.000
1.000	2.000	1.000
0.000	5.000	5.000
3.000	5.000	2.000
2.000	4.000	2.000
0.000	4.000	4.000

What your data sheet should look like

NOW, create a difference score for the grades. Steps 1 and 2 remain the same, step 3 is just a little different – it should look like this:



Creating the “grades” difference score – quarter 2 subtracted from quarter 3

L	M	N	O	P	Q
Q2	Q3	pretest total	posttest total	quiz dif	grade diff
81	88	1.000	4.000	3.000	7.000
79	87	2.000	5.000	3.000	8.000
87	86	1.000	2.000	1.000	-1.000
62	84	0.000	5.000	5.000	22.000
80	91	3.000	5.000	2.000	11.000
81	84	2.000	4.000	2.000	3.000
54	78	0.000	4.000	4.000	24.000

What your data sheet should look like

EZAnalyze Tutorial 1

MAKING SENSE OF PAT'S DATA

Obtaining Descriptive Statistics

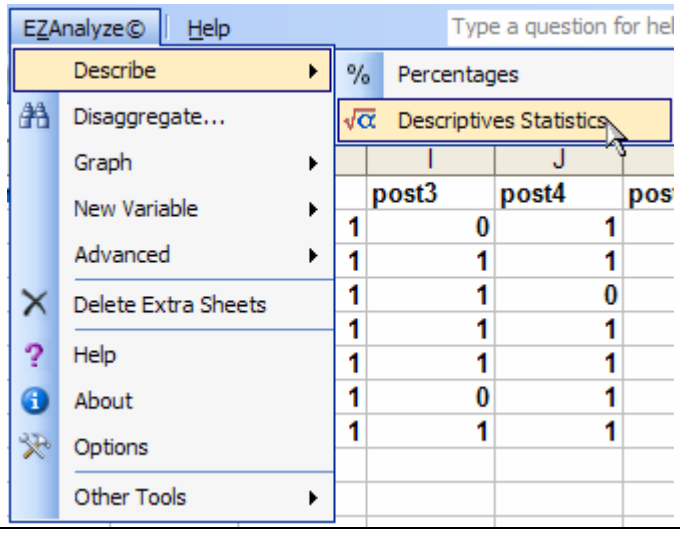
NOW that we have put all of Pat's student data into an Excel data sheet and created summary and difference scores for each student, it is time to turn to the questions Pat would like to answer with the data:

1. **How much did student scores increase from pretest to posttest?**
2. **How much did student grades increase from quarter two to quarter three?**

To answer these questions, we need to be able to say something about the group as a whole – we want to be able to answer these questions with as few numbers as possible to summarize the impact of the intervention.

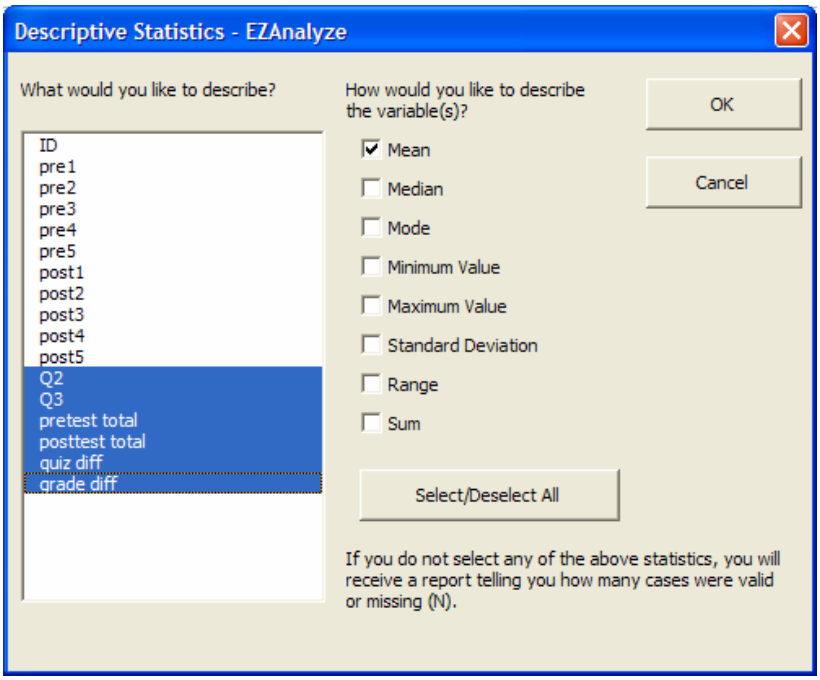
In this case, an appropriate statistic to use to answer both of these questions is the *Mean* or *Average* score for the entire group. EZAnalyze makes computing the mean and other descriptive statistics simple with the DESCRIBE FUNCTION. To use the Describe function to obtain mean scores, follow these steps:

STEP 1



Go to the EZAnalyze Menu and select "Describe", then "Descriptive Statistics"

STEP 2



Highlight the "Q2" – "grade diff" variables by clicking on each one. Make sure "Mean" is selected as the method for describing the variables, and click OK.

EZAnalyze Tutorial 1

CONGRATULATIONS! You have created your first EZAnalyze Results Report! You will notice that a new "sheet" with the results of your data analysis has been created for you – this is how EZAnalyze provides your results.

	A	B	C	D	E	F	G
1	*EZAnalyze Results Report - Descriptive Statistics*						
2							
3		Q2	Q3	pretest total	posttest total	quiz diff	grade diff
4	N Valid:	7.000	7.000	7.000	7.000	7.000	7.000
5	N Missing	0.000	0.000	0.000	0.000	0.000	0.000
6	Mean:	74.857	85.429	1.286	4.143	2.857	10.571

The EZAnalyze Results Report for Descriptive Statistics

As you can see, each column contains one of the variables you selected for analysis, while the rows contain some information about your data and the statistic you chose for analysis – in this case, the *Mean*.

N Valid: This indicates the number of responses for each variable that contained usable data.

N Missing: This indicates the number of responses for each variable that were blank – if you did not have the Q3 grade for a student, you would know it by looking here.

INTERPRETING DESCRIPTIVE STATISTICS
Answering Pat's Questions

	A	B	C	D	E	F	G
1	*EZAnalyze Results Report - Descriptive Statistics*						
2							
3		Q2	Q3	pretest total	posttest total	quiz diff	grade diff
4	N Valid:	7.000	7.000	7.000	7.000	7.000	7.000
5	N Missing	0.000	0.000	0.000	0.000	0.000	0.000
6	Mean:	74.857	85.429	1.286	4.143	2.857	10.571

The EZAnalyze Results Report for Descriptive Statistics

All of the information we need to help Pat answer the questions are located in the EZAnalyze Results Report.

To answer Pat's first question, "**How much did student scores increase from pretest to posttest?**", we can use the Mean that is reported for the "quiz diff" variable:

ANSWER: *Student scores increased 2.86 points on average from the pretest quiz to the posttest quiz.*

We can see that the mean for the pretest total was 1.29, and the mean for the posttest total was 4.14; the amount of increase is indicated by the "quiz diff" mean, which is 2.86.

To answer Pat's second question, "**How much did student grades increase from quarter two to quarter three?**", we can use the Mean that is reported for the "grade diff" variable:

ANSWER: *Student grades increased 10.57 points on average from quarter one to quarter two.*

We can see that the mean for Q2 was 74.86 and the mean for Q3 was 85.43; the amount of increase is indicated by the "grade diff" mean, which is 10.57.

DOES IT APPEAR AS THOUGH PAT'S INTERVENTION IS EFFECTIVE?

YES!

EZAnalyze Tutorial 2

Academic Achievement in Pat's School

In this tutorial, Pat has distributed a survey to a group of high school students. Pat hopes that the results of this survey, coupled with the students' grades, can answer several questions:

1. What are the characteristics of people who responded to the survey?
2. How connected do students feel to various school staff members?
3. What do students think about their classes?
4. What is the "achievement gap" (the academic performance of each ethnic group) in this high school?
5. What is the relationship between feeling connected to school staff and academic achievement?
6. How many students from each ethnic group are there in this sample?
7. Do male and female students have similar academic achievement?

On the next page, you will see a copy of the survey that Pat distributed to students in the high school. Pat received 108 surveys back from students, and entered the data into an Excel data sheet named "EZAsample_data.xls." This file is available for free download from www.ezanalyze.com.

Specific Contents of Tutorial 2

Pat's Student Survey

Demographic Information

1. **Are you male or female**? [gender] Male=0 Female=1
2. **Ethnicity [ethnicity]**: Black=1 Latino/Hispanic=2 White=3 Native American=4 Asian=5 Other=6 multiethnic=7
3. **What grade are you in now? [grade level]**
   9 10 11 12

4. **Classes**
   a. **My classes are interesting to me: [classes a]** none of them=0 some of them=1 all of them=2
   b. **My classes are about things that are important to me: [classes b]**
      none of them =0 some of them=1 all of them=2
   c. **I have friends in my classes: [classes c]** none of them=0 some of them=1 all of them=2
   d. **My classes are the right size: [classes d]** none of them=0 some of them=1 all of them=2

5. **Since 6th grade have you participated in any of the following?** (circle all that apply)
   Gifted education [gifted class] 0=no 1=yes
   Special Education[special ed. class] 0=no 1=yes

6. **How connected do you feel to the following people at HS?**

a.	**Other students in my grade**: [connect a]	not connected=1	somewhat connected=2	very connected=3
b.	**Students in other grades**: [connect b]	not connected=1	somewhat connected=2	very connected=3
a.	**Teachers**: [connect c]	not connected=1	somewhat connected=2	very connected=3
b.	**Dean(s)**: [connect d]	not connected=1	somewhat connected=2	very connected=3
c.	**House Administrators**: [connect e]	not connected=1	somewhat connected=2	very connected=3
d.	**Guidance Counselor(s)**: [connect f]	not connected=1	somewhat connected=2	very connected=3
e.	**Stud. Assistance Counselor(s)**: [connect g]	not connected=1	somewhat connected=2	very connected=3
f.	**School Psychologist**: [connect h]	not connected=1	somewhat connected=2	very connected=3
g.	**Secretaries**: [connect i]	not connected=1	somewhat connected=2	very connected=3
h.	**Custodial staff**: [connect j]	not connected=1	somewhat connected=2	very connected=3
i.	**Cafeteria staff**: [connect k]	not connected=1	somewhat connected=2	very connected=3

Each of the questions above is a variable in Pat's data sheet. The variable name is in [brackets], *and the number representing each possible student response is also indicated. For example, a student who felt "very connected" to teachers would have a "3" on the data sheet for the "connect c" variable.*

EZAnalyze Tutorial 2

DESCRIBING THE SAMPLE

Using the EZAnalyze PERCENTAGES Function

The first thing we need to help Pat do is describe the sample. Pat has 108 students in his sample; obtaining some percentages will help us get acquainted with the data and students in Pat's school – and answer Pat's first question: *What are the characteristics of people who responded to the survey?*

To do this, we will use EZAnalyze to answer the following questions:

1. How many males and females responded to the survey?
2. How many students from each grade level responded to the survey?
3. How many students took special education and gifted classes?

THE PERCENTAGES FUNCTION

When to use PERCENTAGES: Use the percentages function when you have questions such as the ones listed above – in general, you use percentages to describe what your sample looks like.

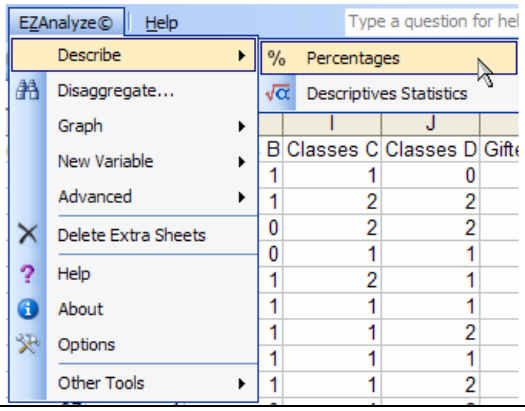
Why to use PERCENTAGES: You use the percentages function to gain an initial understanding of your data. The Percentages function of EZAnalyze allows you to *describe* your sample. This is intentionally kept separate from the EZAnalyze DESCRIPTIVE STATISTICS function because you use PERCENTAGES to describe your *sample*, and DESCRIPTIVE STATISTICS to describe your *results*.

How to use PERCENTAGES: You access the PERCENTAGES function by going to DESCRIBE on the EZAnalyze main menu – a walkthrough is on the next page.

HOW TO OBTAIN PERCENTAGES

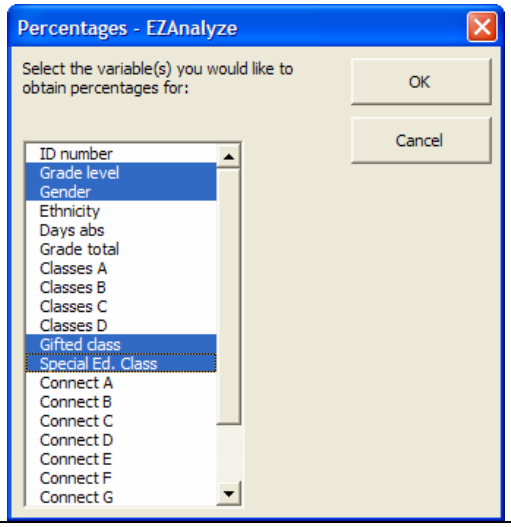
Using the EZAnalyze PERCENTAGES Function

STEP 1



Go to the EZAnalyze Menu and select "Describe", then select "Percentages" from the sub-menu

STEP 2

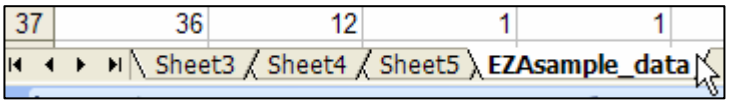


Highlight the "Grade level, Gender, Gifted class, and Special ed. class" variables by clicking on each one. Click OK when you have finished.

EZAnalyze Tutorial 2

When you click OK, you should see the EZAnalyze Results Report for Percentages appear on a new data sheet.

NOTE: to switch back to the data sheet, locate the "EZAsample_data" tab in the lower left hand side of your screen!



INTERPRETING THE EZANALYZE *PERCENTAGES* RESULTS REPORT

	A	B	C	D	E
1	***EZAnalyze Results Report - Percentages***				
2					
3	***Grade level***				
4		**Frequency**	**Percent**	**Valid Percent**	**Cumulative Percent**
5	9	41.000	37.963	37.963	37.963
6	10	28.000	25.926	25.926	63.889
7	11	25.000	23.148	23.148	87.037
8	12	14.000	12.963	12.963	100.000
9					
10	***Gender***				
11		**Frequency**	**Percent**	**Valid Percent**	**Cumulative Percent**
12	0	80.000	74.074	74.074	74.074
13	1	28.000	25.926	25.926	100.000
14					
15	***Gifted class***				
16		**Frequency**	**Percent**	**Valid Percent**	**Cumulative Percent**
17	MISSING	1.000	0.926		
18	0	98.000	90.741	91.589	91.589
19	1	9.000	8.333	8.411	100.000
20					
21	***Special Ed. Class***				
22		**Frequency**	**Percent**	**Valid Percent**	**Cumulative Percent**
23	0	83.000	76.852	76.852	76.852
24	1	25.000	23.148	23.148	100.000

41 ninth grade students responded to the survey (this is the **FREQUENCY**)

Just under 8.3% of students have taken a gifted class (this is the **PERCENT)**

Just over 23% of students have taken a special education class.

What your Results Report should look like

Elements of the EZAnalyze *PERCENTAGES* Results Report

VARIABLES selected for analysis are always listed in the first column (column A). The numbers under the variable name are the responses listed in the data set.

FREQUENCY is the actual number of people for each response.

VALID PERCENT is the percent of the sample who had *valid* responses for the variable. This is only important if there are *missing values* – people who did not answer a question. There is a single *missing value* for the gifted class variable as an example.

CUMULATIVE PERCENT is the percent of people who had *at least* the response given. For example, 73% of the sample responded with a "3" or lower for the "Ethnicity" variable. In other words, 73% of the sample indicated they were Black (1), Latino/Hispanic (2), or White (3).

EZAnalyze Tutorial 2

ANSWERING PAT'S QUESTION

What are the characteristics of people who responded to the survey?

	A	B	C	D	E
1	***EZAnalyze Results Report - Percentages***				
2					
3	***Grade level***				
4		**Frequency**	**Percent**	**Valid Percent**	**Cumulative Percent**
5	9	41.000	37.963	37.963	37.963
6	10	28.000	25.926	25.926	63.889
7	11	25.000	23.148	23.148	87.037
8	12	14.000	12.963	12.963	100.000
9					
10	***Gender***				
11		**Frequency**	**Percent**	**Valid Percent**	**Cumulative Percent**
12	0	80.000	74.074	74.074	74.074
13	1	28.000	25.926	25.926	100.000
14					
15	***Gifted class***				
16		**Frequency**	**Percent**	**Valid Percent**	**Cumulative Percent**
17	MISSING	1.000	0.926		
18	0	98.000	90.741	91.589	91.589
19	1	9.000	8.333	8.411	100.000
20					
21	***Special Ed. Class***				
22		**Frequency**	**Percent**	**Valid Percent**	**Cumulative Percent**
23	0	83.000	76.852	76.852	76.852
24	1	25.000	23.148	23.148	100.000

To answer this question, we will use the Percentages EZAnalyze Results Report to obtain answers to three sub-questions:

1. How many males and females responded to the survey?

This question can be answered by looking at the "Gender" variable Frequency and Percent

ANSWER – 80 males (74%) and 28 females (26%)

2. How many students from each grade level responded to the survey?

This question can be answered by looking at the "Grade level" variable Frequency and Percent

ANSWER – 41 ninth grade students (38%), 28 tenth grade students (26%), 25 eleventh grade students (23%), and 14 twelfth grade students (13%).

3. How many students took special education and gifted classes?

This question can be answered by looking at the "Gifted" and "Special ed. class" variables' Frequency and Percent

ANSWER – 9 students took a gifted class (8.3%), and 25 a Special Education class (23%)

EZAnalyze Tutorial 2

DESCRIBING YOUR RESULTS

Using the EZAnalyze DESCRIPTIVE STATISTICS Function

To answer Pat's second and third questions, *How connected do students feel to various school community members?* and *What do students think about their classes?*, we will use the DESCRIPTIVE STATISTICS function of EZAnalyze. This function allows you to obtain some results data to answer questions with your data.

In this example, we want to know how connected students feel to various members of the school community – namely, other students and staff, and what students think about their classes. These questions are represented in the data set as variables "Connect A" through "Connect K" for the connectedness questions, and "Classes A" through "Classes D." To answer this question, we will obtain an average or "mean" score for each of these variables, and use the results report to interpret what the data are saying.

THE DESCRIPTIVE STATISTICS FUNCTION

When to use DESCRIPTIVE STATISTICS: Use the descriptive statistics function when you want to obtain some results data to answer a question you have about a set of variables – in this case, we are interested in knowing which variable has the highest mean score. If you wanted to know what the average grade total was for your sample - or any of the descriptive statistics described below – you can use this function

Why to use DESCRIPTIVE STATISTICS: You use the descriptive statistics function to generally describe the results contained your data.

How to use DESCRIPTIVE STATISTICS: You access the DESCRIPTIVE STATISISTICS function by going to DESCRIBE on the EZAnalyze main menu – a walkthrough is on the next page.

- **Mean** – *The "average" number; calculated by adding up all of the scores, and dividing by the total number scores*
- **Median** –*The "middle" number; calculated by ordering all the numbers from low to high, then selecting the middle number*
- **Mode** – *The most frequently occurring number; calculated by determining the number that occurred most often*
- **Minimum Value** – *The "lowest" number – calculated by ordering all the numbers from low to high, then selecting the lowest number*
- **Maximum Value** – *The "highest" number – calculated by ordering all the numbers from low to high, then selecting the highest number*
- **Standard Deviation** – *The number indicating how far, on average, scores were different from the average or "mean" score*
- **Range** – *The number indicating how far it is from the lowest score to the highest score. It is the max value minus the min value*
- **Sum** – *The value of all of the scores in the selected variable added together*

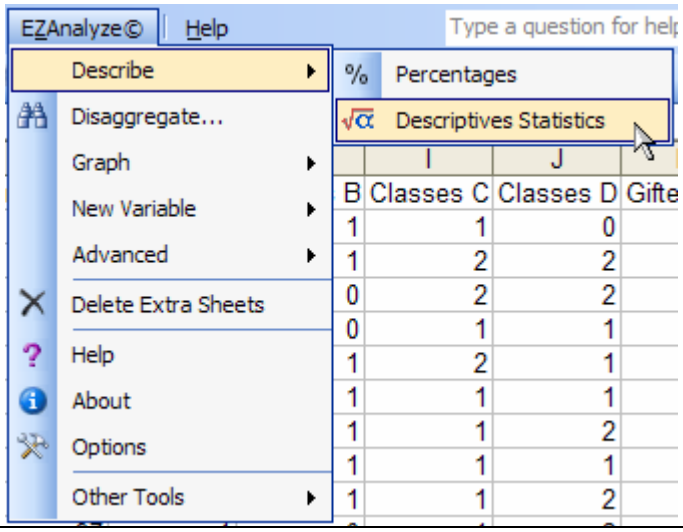
Definitions of the EZAnalyze Descriptive Statistics

EZAnalyze Tutorial 2

HOW TO OBTAIN DESCRIPTIVE STATISTICS

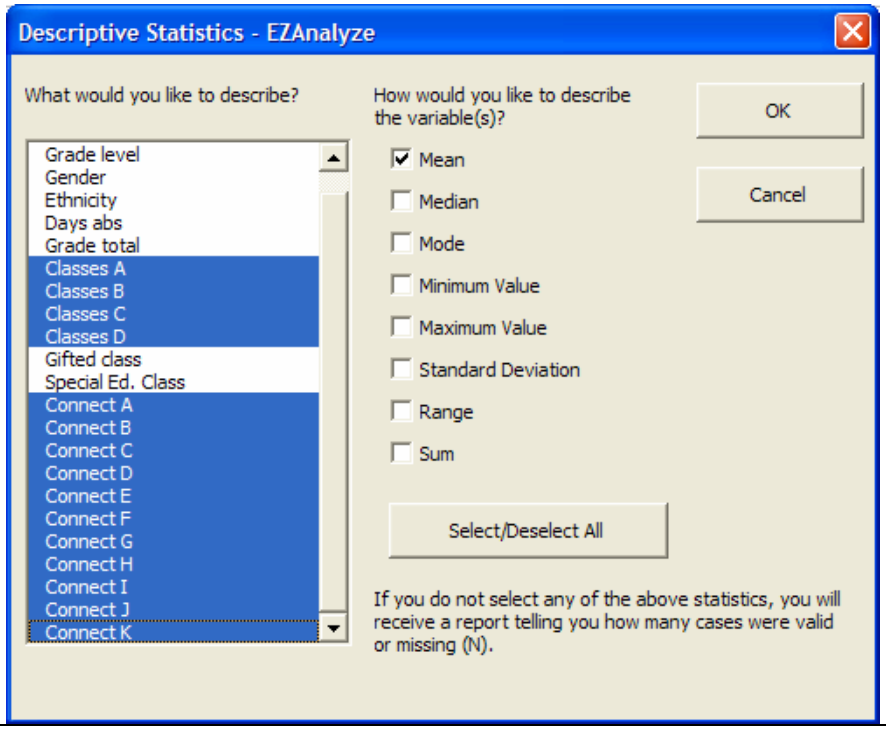
Using the EZAnalyze DESCRIPTIVE STATISTICS Function

STEP 1



Go to the EZAnalyze Menu and select "Describe", then select "Descriptive Statistics" from the sub-menu

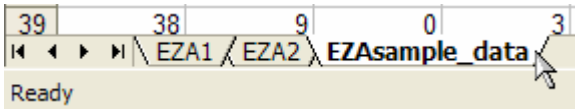
STEP 2



Highlight the variables "Classes A" to " Classes D" and "Connect A" to "Connect K" by clicking each one. Select "Mean" as the method for describing your variables, and click OK.

When you click OK, you should see the EZAnalyze Results Report for Descriptive statistics appear on a new data sheet

NOTE: to switch back to the data sheet, locate the "EZAsample_data" tab in the lower left hand side of your screen!



EZAnalyze Tutorial 2

INTERPRETING THE EZANALYZE *DESCRIPTIVE STATISTICS* RESULTS REPORT

	A	B	C	D	E	F	G	H	I
1	***EZAnalyze Results Report - Descriptive Statistics***								
2									
3		**Classes A**	**Classes B**	**Classes C**	**Classes D**	**Connect A**	**Connect B**	**Connect C**	**Connect D**
4	**N Valid:**	108.000	108.000	108.000	108.000	108.000	108.000	108.000	108.000
5	**N Missing**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
6	**Mean:**	0.963	0.907	1.361	1.324	1.611	1.981	2.204	2.593

	A	J	K	L	M	N	O	P
1	***EZAnalyz***							
2								
3		**Connect E**	**Connect F**	**Connect G**	**Connect H**	**Connect I**	**Connect J**	**Connect K**
4	**N Valid:**	108.000	108.000	108.000	108.000	108.000	108.000	108.000
5	**N Missing**	0.000	0.000	0.000	0.000	0.000	0.000	0.000
6	**Mean:**	2.519	2.093	2.630	2.741	2.602	2.759	2.787

What your results report should look like (modified to fit this page)

Connect A has the lowest *Connect* score — Connect K has the highest *Connect* score

Classes B has the lowest *Classes* score — Classes C has the highest *Classes* score

Elements of the EZAnalyze *DESCRIPTIVE STATISTICS* Results Report

VARIABLES selected for analysis are always listed in the columns (columns B through P here).

N VALID is the number of valid responses; **N MISSING** is the number of missing responses – adding these together will give you your total sample size.

STATISTICS selected for reporting are listed under the "N Missing" – "N Valid" and "N Missing" are always reported. Here, we selected "Mean"

The remaining rows of the EZAnalyze Descriptive Statistics Results Report will contain the statistics you selected for analysis under "Statistic Reported."

ANSWERING PAT'S QUESTIONS

How connected do students feel to various school community members?
What do students think about their classes?

What do students think about their classes?
This question can be answered by looking at the Means for each of the "Class" variables

ANSWERS
- They have friends in many of their classes (Class C)
- Their classes are usually the right size (Class D)
- Some of their classes are interesting to them (Class A)
- Some of their classes are important to them (Class B)

What is the school community member students feel most connected to?
This question can be answered by looking at the Means for each of the variables

ANSWER – Cafeteria staff (Connect K)

What is the school community member students feel least connected to?
This question can be answered by looking at the Means for each of the variables

ANSWER – Other students in their grade (Connect A)

EXTRA! Want to try graphing these results? Go to the *Graph* menu in EZAnalyze and create a *Multiple Variable* Graph using these same variables

SUMMARY: Students overall feel more connected to staff members of the school community than they do to other students. Students feel most connected to the cafeteria staff, followed by custodial staff, the school psychologist, the student assistance counselor, secretaries, the dean, house administrators, and teachers. Students felt least connected to other students in their grade, followed by students in other grades and guidance counselors (this is fake data, remember that!).

EZAnalyze Tutorial 2

DISAGGREGATING DATA

Using the EZAnalyze DISAGGREGATE Function

To answer Pat's second question, *What is the "achievement gap" (the academic performance of each ethnic group) in this high school?*, we will need to *Disaggregate* the academic achievement data by ethnicity.

What this means is that we will find out what the "average" (or "mean") grade is for each of the ethnic groups in the sample – we will break apart the academic achievement data we have by students' ethnicity.

THE DISAGGREGATE FUNCTION

When to use DISAGGREGATE: Use the disaggregate function when you want to see how different subgroups score on a variable.

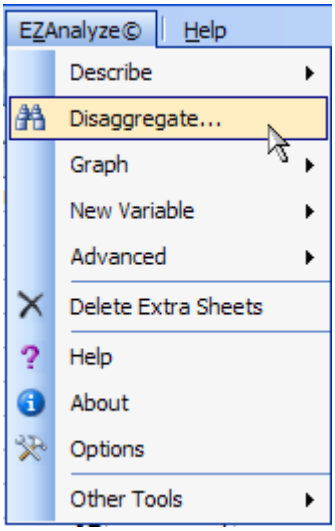
Why to use DISAGGREGATE: You use the disaggregate function of EZAnalyze to sort your analysis of one variable into categories. If you want to see how differently males and females score on a variable like grades, or how different ethnic groups perform academically, you use the disaggregate function to break the data apart into groups for you.

How to use DISAGGREGATE: You access the DISAGGREGATE function by going to DISAGGREGATE on the EZAnalyze main menu. To use the disaggregate function, you need to have two variables in mind – a *dependent variable* and a *categorical variable.* The dependent variable is the variable you are interested in breaking apart, while your categorical variable is the variable that tells EZAnalyze what the groups are. For example, "grades" is an example of a dependent variable, and "gender" and "ethnicity" are examples of categorical variables. A walkthrough is below for answering Pat's question – you will learn to disaggregate "grades" by "ethnicity."

HOW TO DISAGGREGATE DATA

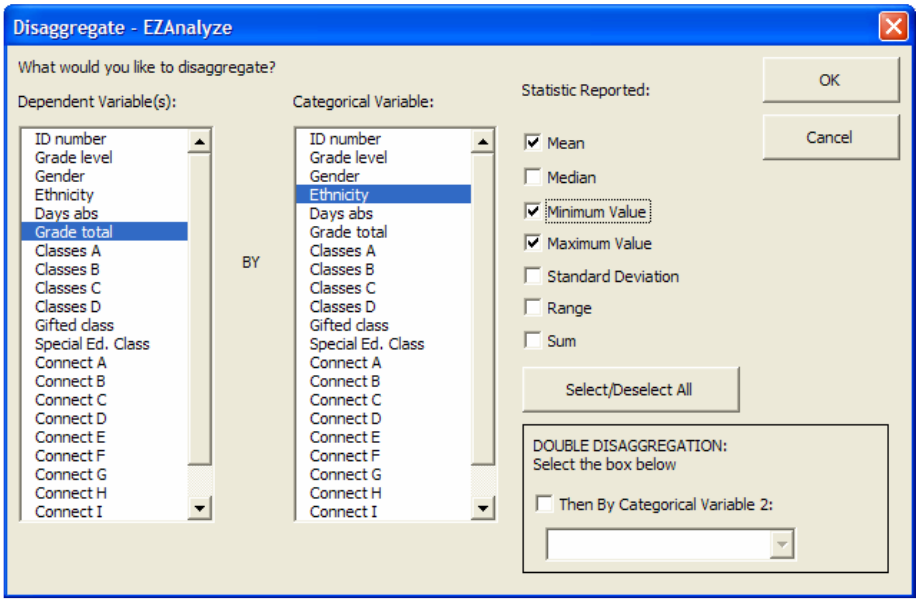
Using the EZAnalyze DISAGGREGATE Function

STEP 1



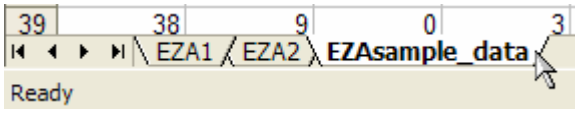
Go to the EZAnalyze Menu and select "Disaggregate"

STEP 2



Select "Grade total" under Dependent Variable, and "Ethnicity" under Categorical Variable. Check off "Mean" under Statistic Reported, and check off "Minimum Value" and "Maximum Value". Click OK when you have finished.

When you click OK, you should see the EZAnalyze Results Report for Disaggregate appear on a new data sheet.



NOTE: to switch back to the data sheet, locate the "EZAsample_data" tab in the lower left hand side of your screen!

EZAnalyze Tutorial 2

INTERPRETING THE EZANALYZE *DISAGGREGATE* RESULTS REPORT

	A	B	C	D	E	F	G
1	***EZAnalyze Results Report - Grade total Disaggregated by Ethnicity***						
2							
3	**Ethnicity**		**N**	**Minimum**	**Maximum**	**Mean**	
4	1	**Grade total**	16.000	63.070	89.210	79.636	
5	2	**Grade total**	13.000	62.150	89.870	76.949	
6	3	**Grade total**	50.000	75.720	99.540	91.878	
7	5	**Grade total**	10.000	76.840	97.250	90.285	
8	6	**Grade total**	5.000	75.220	89.640	81.998	
9	7	**Grade total**	14.000	79.070	88.950	84.603	
10							

What your Results Report should look like

The "Latino/Hispanic" group had the lowest average grade

The "White" group had the highest average grade

Elements of the EZAnalyze *DISAGGREGATE* Results Report

The CATEGORICAL VARIABLE selected for analysis is always listed in the first column (column A). The numbers under the categorical variable name are the responses listed in the data set.

The DEPENDENT VARIABLE(S) selected for analysis will always be in the second column (column B). If we had more than one variable, they would be listed underneath the first dependent variable for each level of the categorical variable.

The N (column C) is the number of valid responses. If there are missing values, this number will tell you how many valid cases there are for each group.

The remaining columns of the EZAnalyze Disaggregate Results Report will contain the statistics you selected for analysis under "Statistic Reported."

- **Mean** – *The "average" number; calculated by adding up all of the scores, and dividing by the total number scores*
- **Median** –*The "middle" number; calculated by ordering all the numbers from low to high, then selecting the middle number*
- **Minimum Value** – *The "lowest" number – calculated by ordering all the numbers from low to high, then selecting the lowest number*
- **Maximum Value** – *The "highest" number – calculated by ordering all the numbers from low to high, then selecting the highest number*
- **Standard Deviation** – *The number indicating how far, on average, scores were different from the average or "mean" score*
- **Range** – *The number indicating how far it is from the lowest score to the highest score. It is the max value minus the min value*
- **Sum** – *The value of all of the scores in the selected variable added together*

Definitions of the EZAnalyze Descriptive Statistics

Extra Practice – You can create a graph of these results by creating a *Disaggregation Graph* – go into the *Graph* menu of EZAnalyze and select *Disaggregation Graph*; then simply enter the dependent and categorical variables just as you did for the analysis above!

EZAnalyze Tutorial 2

ANSWERING PAT'S QUESTION

What is the "achievement gap" (the academic performance of each ethnic group) in this school?

	A	B	C	D	E	F	G
1	***EZAnalyze Results Report - Grade total Disaggregated by Ethnicity***						
2							
3	**Ethnicity**		**N**	**Minimum**	**Maximum**	**Mean**	
4	**1**	**Grade total**	16.000	63.070	89.210	79.636	
5	**2**	**Grade total**	13.000	62.150	89.870	76.949	
6	**3**	**Grade total**	50.000	75.720	99.540	91.878	
7	**5**	**Grade total**	10.000	76.840	97.250	90.285	
8	**6**	**Grade total**	5.000	75.220	89.640	81.998	
9	**7**	**Grade total**	14.000	79.070	88.950	84.603	
10							

To answer this question, we will use the Disaggregate EZAnalyze Results Report to obtain the average "grade total" for each ethnic group

What is the average grade for the "Black" ethnicity? *ANSWER* – 79.6
This question can be answered by looking at the Mean for "grade total" for "ethnicity"=1

What is the average grade for the "Latino/Hispanic" ethnicity? *ANSWER* – 76.9
This question can be answered by looking at the Mean for "grade total" for "ethnicity"=2

What is the average grade for the "White" ethnicity? *ANSWER* – 91.9
This question can be answered by looking at the Mean for "grade total" for "ethnicity"=3

What is the average grade for the "Asian" ethnicity? *ANSWER* – 90.3
This question can be answered by looking at the Mean for "grade total" for "ethnicity"=5

What is the average grade for the "Multiethnic" ethnicity? *ANSWER* – 84.6
This question can be answered by looking at the Mean for "grade total" for "ethnicity"=7

SUMMARY: An achievement gap is evident in Pat's high school. Students who characterized themselves as "White" or "Asian" have, on average, grades above 90. Students who characterized themselves as "Other" or "Multiethnic" have, on average, grades in the low to mid 80's, while students who characterize themselves as "Black" or "Latino/Hispanic" have, on average, grades in the mid to upper 70's.

EZAnalyze Tutorial 2

DESCRIBING RELATIONSHIPS

Using the EZAnalyze CORRELATE Function

To answer Pat's fifth question, *What is the relationship between feeling connected to school community members and academic achievement?*, we will use the CORRELATE function of EZAnalyze on the "Advanced" menu. This function allows you to obtain a *correlation coefficient* – a number used to describe the relationship between two variables.

In this example, we want to know how feeling connected to school community members is related to academic achievement. To answer this question, we will first create a total "connectedness" score for each student, then use the EZAnalyze CORRELATE function to obtain a correlation coefficient.

THE CORRELATE FUNCTION

When to use CORRELATE: Use the CORRELATE function when you want to see how two variables are related to each other. Other examples of places to use the correlate function in education would be to determine the relationship between the number of days absent and academic achievement, or the relationship between the number behavioral referrals and academic achievement.

Why to use CORRELATE: You use the correlate function to obtain a single number that describes the relationship between two variables.

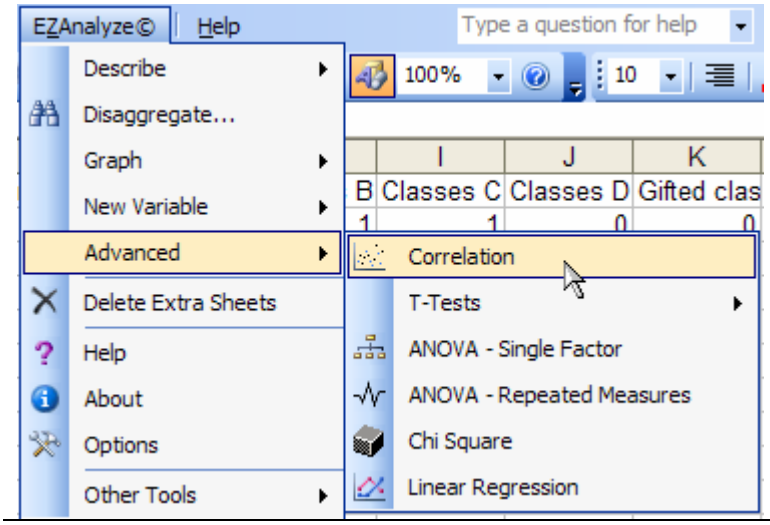
How to use CORRELATE: You access the CORRELATE function by going to CORRELATE on the EZAnalyze Advanced menu. The result of the correlate function is a *correlation coefficient.* This correlation coefficient, a number, ranges from -1 to +1. The *direction* of the relationship is indicated by the sign (- or +), while the *strength* of the relationship is indicated by the number. Correlation coefficients that are close to zero indicate a small relationship, while numbers that are close to one indicate a strong relationship. For example, a correlation coefficient of "-.86" would indicate a relatively strong, negative relationship, while a correlation coefficient of ".32" would indicate a relatively weak positive relationship. A walkthrough is below.

NOTE: *Before continuing to the next page, create a new SUMMARY VARIABLE named "connectedness" for all of the "Connect" variables (Connect A – K). If you do not know how to do this already, view the tutorial on page 8.*

HOW TO OBTAIN A CORRELATION COEFFICIENT

Using the EZAnalyze CORRELATE Function

STEP 1

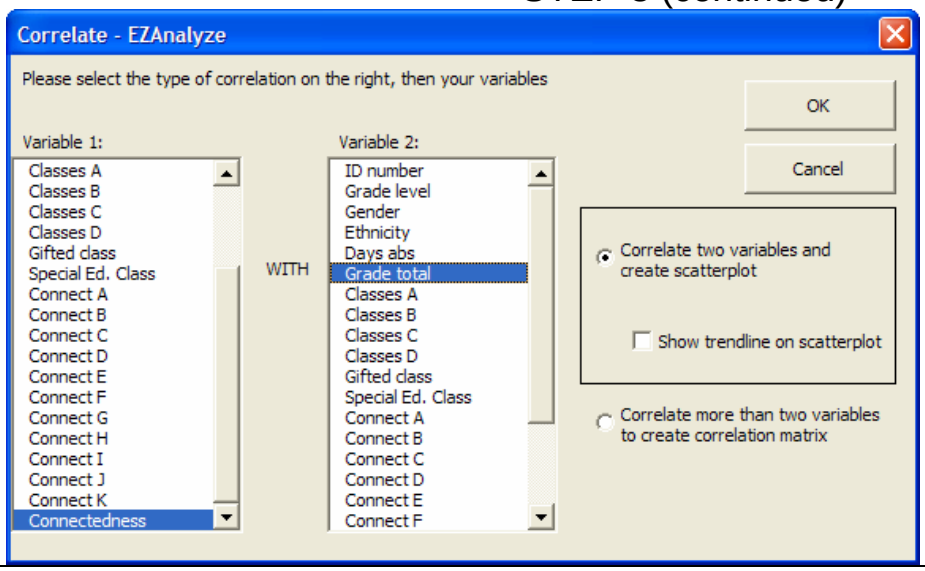


Go to the EZAnalyze Menu and select "Advanced", then select "Correlation" from the sub-menu

NEXT

EZAnalyze Tutorial 2

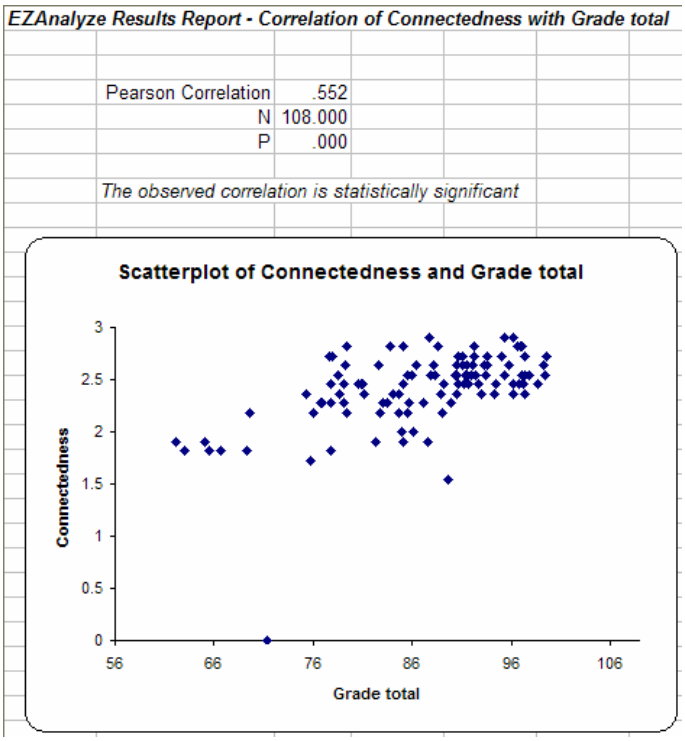
STEP 3 (continued)



Select "connectedness" under the Variable 1 column, and "Grade total" under the Variable 2 column. Click OK when you have finished.

When you click OK, you should see the EZAnalyze Results Report for Correlate appear on a new data sheet.

INTERPRETING THE EZANALYZE *CORRELATE* RESULTS REPORT



What your results report should look like

P tells you if the correlation is statistically significant

The scatterplot provides you with a picture of each person's score. Notice how there was 1 person who had a connectedness score of zero, and a grade of about 72

Elements of the EZAnalyze *CORRELATE* Results Report

VARIABLES selected for analysis are listed in the report title, as well as on the scatterplot

PEARSON CORRELATION is the correlation coefficient.

N is the number of valid responses used to compute the correlation coefficient. If there were missing values, you would find out by looking here.

P is the significance level of the observed correlation.

EZAnalyze Tutorial 2

ANSWERING PAT'S QUESTION

What is the relationship between feeling connected to school community and academic achievement?

EZAnalyze Results Report - Correlation of Connectedness with Grade total						
	Pearson Correlation	.552				
	N	108.000				
	P	.000				
	The observed correlation is statistically significant					

The correlation between "Grade total" and "connectedness" is .55, and is significant at the P < .001 level

---

What is the relationship between feeling connected to school community members and academic achievement?

ANSWER – The correlation coefficient describing the relationship between the connectedness total score and grade is .55

SUMMARY: The relationship between academic achievement, as measured by the students total grade, and "connectedness," as measured by the summed total of all of the *connect* items on the survey, is .55. This indicates a moderate, positive relationship between feeling connected to various school community members and academic achievement – as a student's connected score increases, so too does their academic achievement.

NOTE: When interpreting a correlation coefficient, always remember that CORRELATION DOES NOT IMPLY CAUSATION. Just because two variables are highly correlated does not mean that one causes the other – it only means that they are related. In the example above, we could be tempted to conclude that feeling connected to school *leads to* improved grades. The improved grades could actually be the result of a third variable that we did not measure – such as attendance. ALSO, there is no way to tell with a correlation coefficient which of the two variables causes the other – only that they are related. SO, does feeling connected to school lead to improved grades? OR, do higher grades lead to feeling more connected to school? The methods for determining *causation* are beyond the scope of EZAnalyze's capabilities. HOWEVER, knowing how two variables are related is still valuable information!

EZAnalyze Tutorial 2

DESCRIBING THE SAMPLE WITH A GRAPH

Using the EZAnalyze HISTOGRAM Function

We discussed how to describe a sample with numbers using the PERCENTAGES function early in this tutorial. Now we are going to learn how to describe your sample by using the EZAnalyze HISTOGRAM function to answer Pat's sixth question, *How many students from each ethnic group are there in this sample?* This function allows you to create a graph that shows the *distribution* of scores on a given variable.

In this example, we are going to create a Histogram of the "Ethnicity" variable to see how many students from each ethnic group responded to the survey

THE HISTOGRAM FUNCTION

When to use HISTOGRAM: Use the Histogram function when you want to create a picture of a piece of your data – in this case, we want to "see" how the different grade levels are represented in our sample.

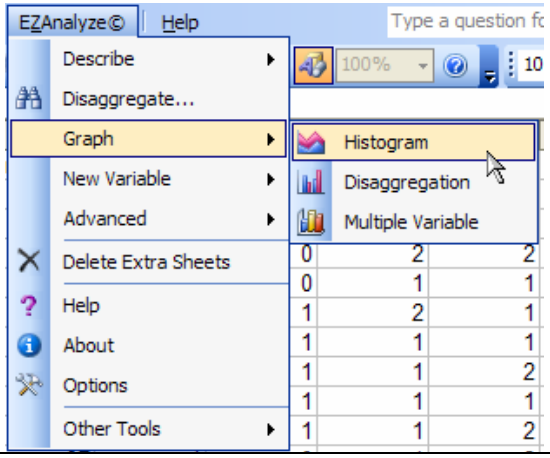
Why to use HISTOGRAM: You use the Histogram function when it is important to create a visual depiction of your data. Graphs are a great tool in general for displaying the results of data analyses in reports that you write, or in presentations.

How to use HISTOGRAM: You access the HISTOGRAM function by going to GRAPHS on the EZAnalyze main menu and select HISTOGRAM. A walkthrough of how to create a Histogram begins below.

HOW TO OBTAIN A HISTOGRAM

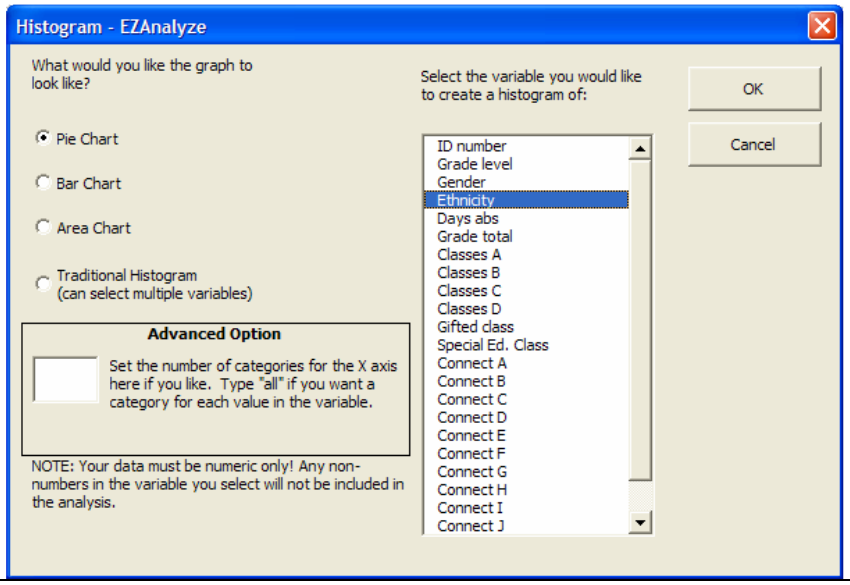
Using the EZAnalyze HISTOGRAM Function

STEP 1



Go to the EZAnalyze Menu and select "Graph", then select "Histogram" from the sub-menu

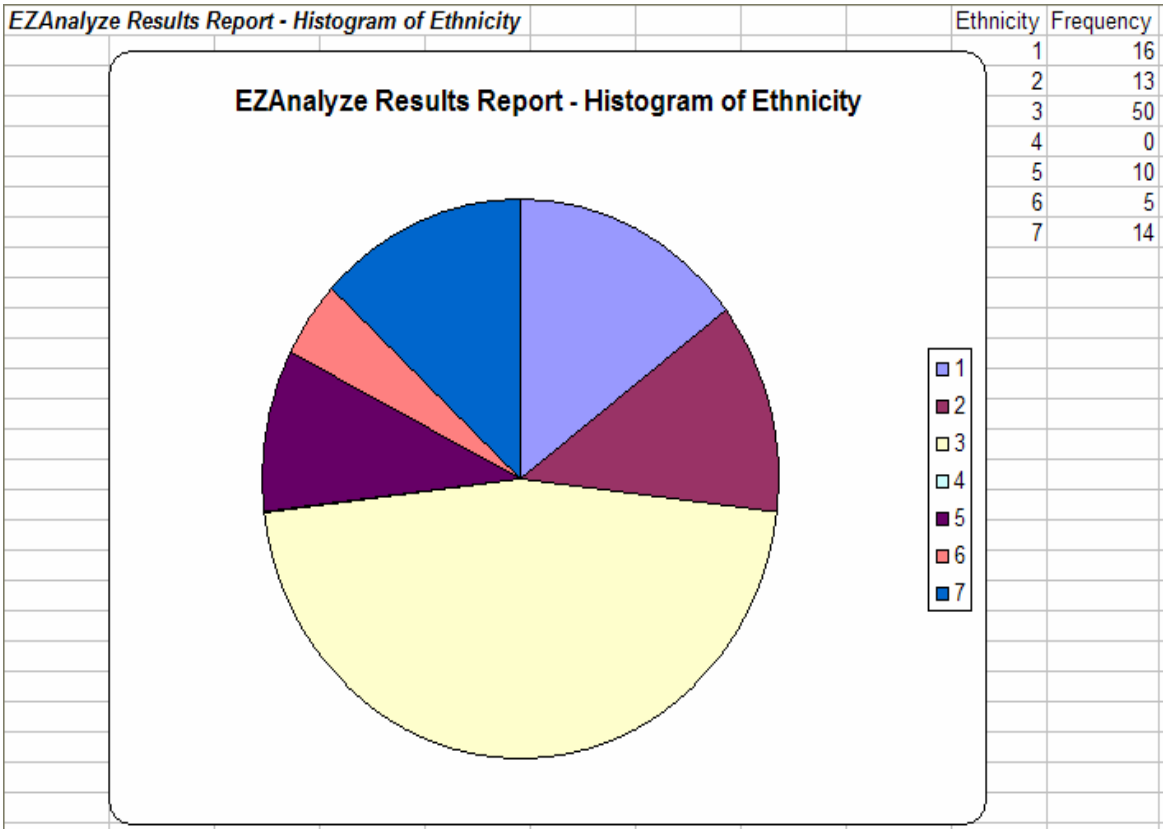
STEP 2



Highlight the "Ethnicity" variable and make sure that "Pie Chart" is selected as the type of graph you want to create. Click OK when you have finished.

EZAnalyze Tutorial 2

INTERPRETING THE EZANALYZE *HISTOGRAM* RESULTS REPORT



What your Results Report should look like

The "legend" on the right contains the numbers that represent the different ethnic groups.

Notice that while there were no "Native American" respondents, there is a "4" in the legend. HOWEVER, the histogram does accurately portray that there are no Native American respondents.

Elements of the EZAnalyze *HISTOGRAM* Results Report

The VARIABLE selected for analysis is always listed in the title of the graph.

The Legend is located on the right side of the Histogram; this is where you will find which numbers correspond to each of the colors.

Before we get into interpreting the Histogram, lets take a few minutes to make our graph a little prettier – and self-explanatory

EZAnalyze Tutorial 2

MODIFYING YOUR HISTOGRAM

EZAnalyze has done the bulk of creating your Histogram for you. You could leave it as is, but in just a few moments, you can create a graph that is truly self-explanatory.

STEP ONE

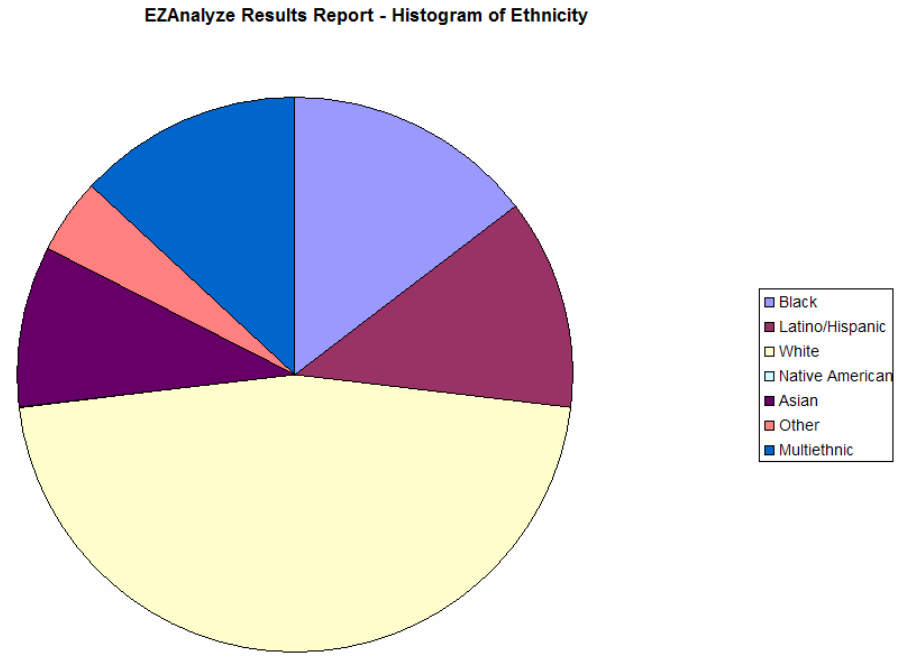
	A	B
1	1	16
2	2	13
3	3	50
4	4	0
5	5	10
6	6	5
7	7	14

Locate the data for your histogram – this is in the upper left corner of your sheet with the graph

STEP TWO

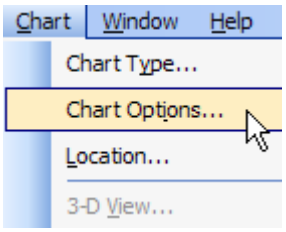
	A	B
1	Black	16
2	Latino/Hispanic	13
3	White	50
4	Native American	0
5	Asian	10
6	Other	5
7	Multiethnic	14

Change the numbers in Column A to the ethnic group names that the numbers represent. Click on the chart sheet when you are done.



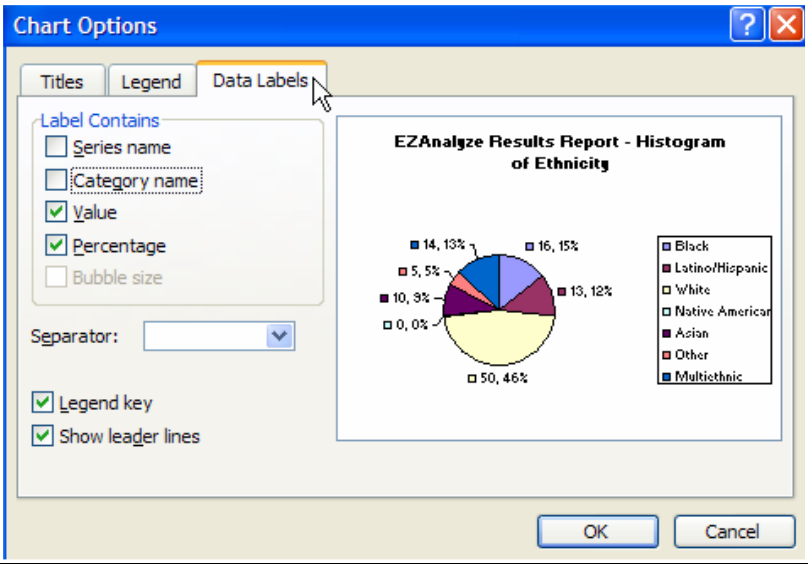
You should see that your Legend has been updated!

STEP THREE



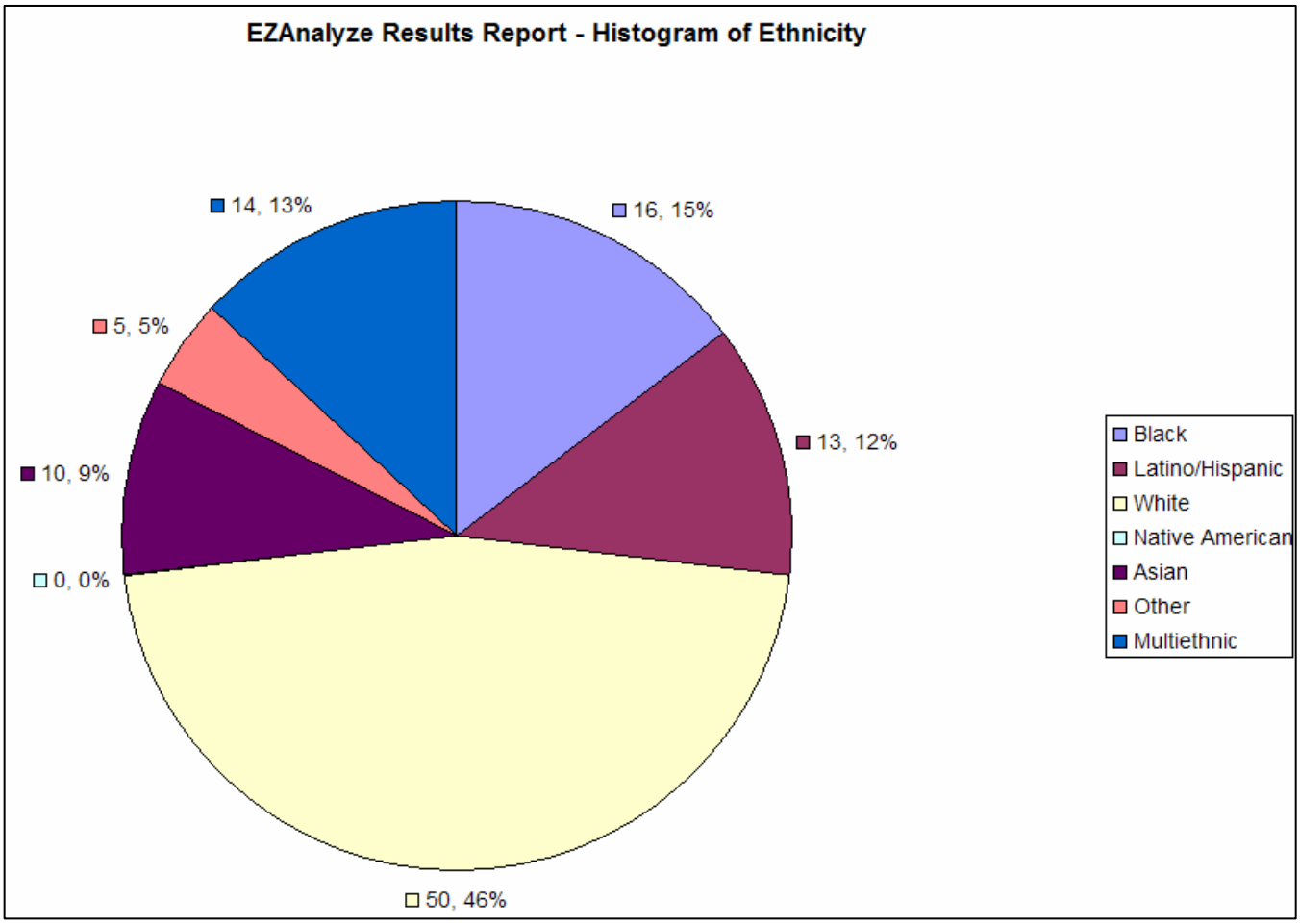
With the Histogram visible, select "Chart Options" from the menu bar

STEP FOUR



In the "Chart options" dialog box that appears, click on the "Data Labels" tab and place a check mark next to "Value" and "Percentage." Click OK when you are done.

EZAnalyze Tutorial 2



You should now see your new and improved Histogram!

NOTE – you can cut and paste this graph into Word PowerPoint!

FOR ADDITIONAL GRAPH MODIFICATIONS, SEE "MODIFYING GRAPHS" IN THE MANUAL

ANSWERING PAT'S QUESTION

How many students from each ethnic group are there in this sample?

To answer this question, we will use the EZAnalyze Histogram to obtain the answers

How many students from each ethnic group are there in the sample?
This question can be answered by looking at the various slices of the Histogram pie and the legend.

ANSWER –16 students (15%) indicated "Black", 13 students (12%) indicated "Latino/Hispanic", 50 (46%) indicated "White", 10 students (9%) indicated "Asian", 5 students (5%) indicated "Other", and 14 students (13%) indicated "Multiethnic" when asked to identify their ethnicity. No students indicated that "Asian" was their ethnicity.

EZAnalyze Tutorial 2

CREATING A DISAGGREGATION GRAPH
Using the EZAnalyze DISAGGREGATION GRAPH Function

We discussed how to disaggregate data with the DISAGGREGATE function earlier in this tutorial. Now we are going to learn how to display disaggregated data using the EZAnalyze DISAGGREGATION GRAPH function to answer Pat's final question, *Do male and female students have similar academic achievement?*
This function allows you to create a graph that displays disaggregated data

In this example, we are going to create a Disaggregation graph to show how males and females performed differently on the "grade total" variable.

THE DISAGGREGATION GRAPH FUNCTION

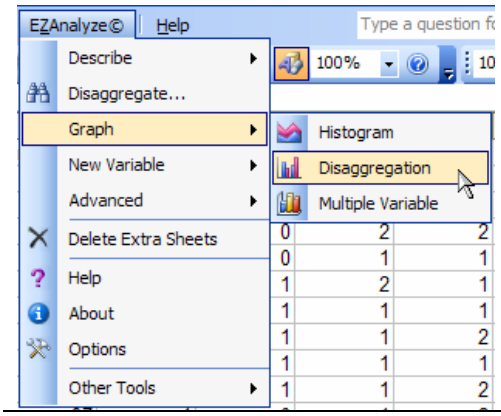
When to use DISAGGREGATION GRAPH: Use the Disaggregation graph function when you want to break apart data on a *dependent variable* by different levels of a *categorical variable*. In this example, we want to break apart "grade total", the dependent variable, by "gender", the categorical variable.

Why to use DISAGGREGATION GRAPH: You use the Disaggregation graph function when you need to create a visual representation of your data analysis. Graphs can be powerful tools for easily communicating the results of your data analyses – particularly disaggregation.

How to use DISAGGREGATION GRAPH: You access the Disaggregation graph function by going to GRAPHS on the EZAnalyze main menu and selecting Disaggregation. A walkthrough of how to create a Disaggregation graph begins below.

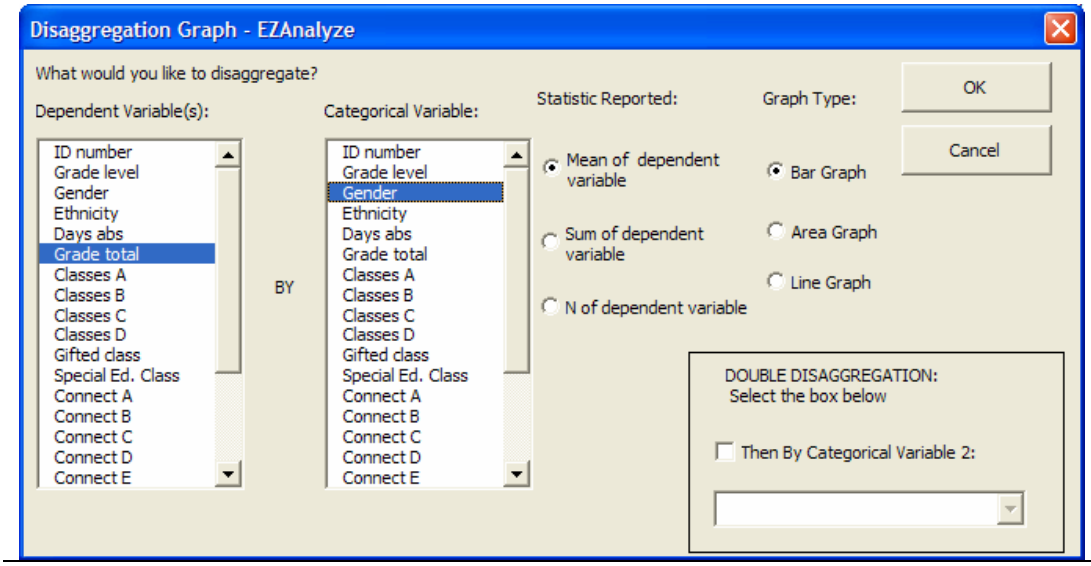
HOW TO OBTAIN A DISAGGREGATION GRAPH
Using the EZAnalyze DISAGGREGATION GRAPH Function

STEP 1



Go to the EZAnalyze Menu and select "Graph", then select "Disaggregation" from the sub-menu

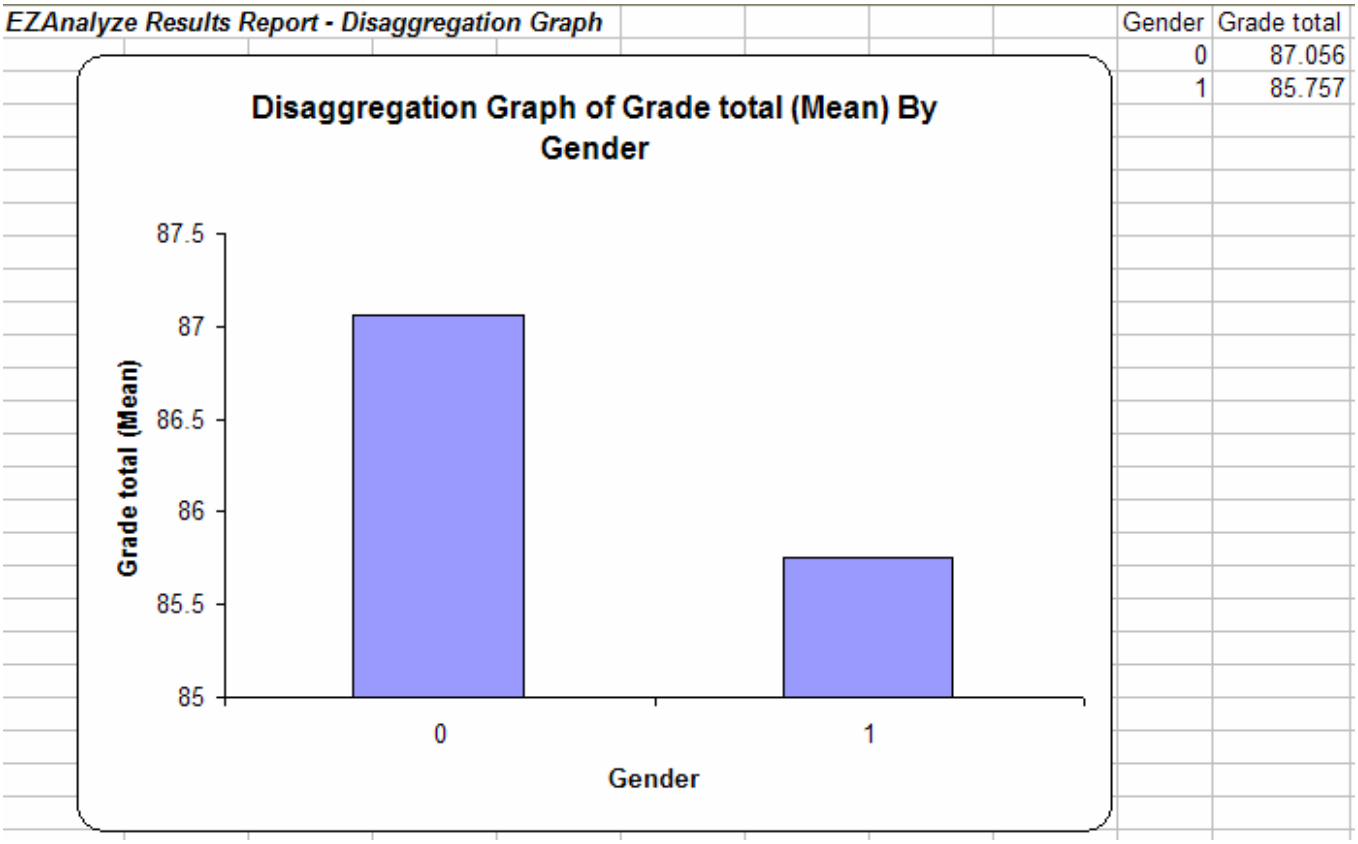
STEP 2



Highlight the "Grade total" variable under Dependent Variable and "Gender" under Categorical Variable. Select "Mean of dependent variable" under Statistic Reported, and "Bar Graph" under Graph Type and click OK

EZAnalyze Tutorial 2

INTERPRETING THE EZANALYZE *DISAGGREGATION GRAPH* RESULTS REPORT



What your Results Report should look like

Each Bar represents one level of the *categorical variable*, while the numbers on the left side of the graph represent the *dependent variable*

Elements of the EZAnalyze *DISAGGREGATION GRAPH* Results Report

The CATEGORICAL VARIABLE selected for analysis is always represented by the bars of the graph, while the **DEPENDENT VARIABLE** is always represented by the numbers on the left side of the graph.

Before we get into interpreting the Disaggregation Graph, take a few minutes to make our graph a little more self-explanatory

---

EZAnalyze Tutorial 2

MODIFYING YOUR DISAGGREGATION GRAPH

To make our Disaggregation Graph "tell the story" a little more effectively, we need to add titles to the graph and change the numbers in our *categorical variable* into the category names

STEP ONE

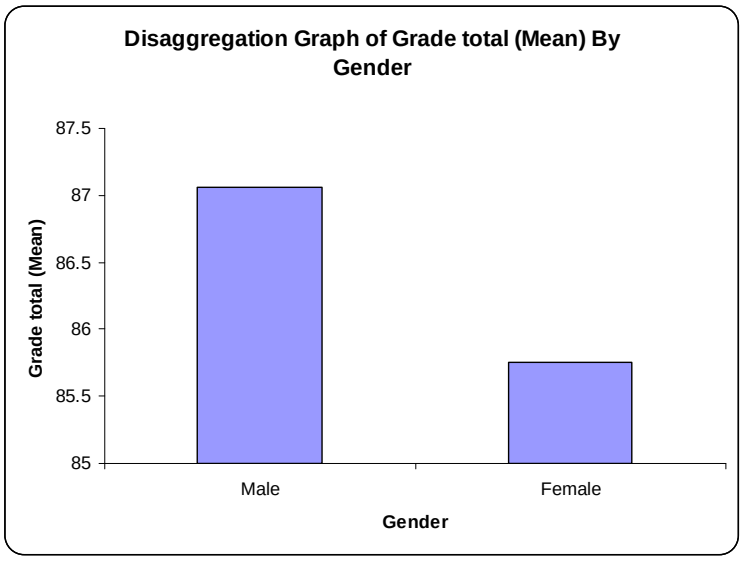
	A	B
1	Gender	Grade total
2	0	87.056
3	1	85.757

Locate the data that are contained in your graph. These are always in the upper left corner of the sheet.

STEP TWO

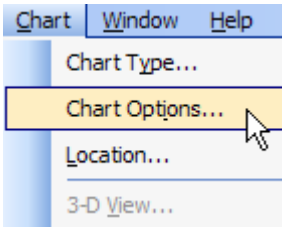
	A	B
1	Gender	Grade total
2	Male	87.056
3	Female	85.757

Change the numbers in Column A to the names "Male" for 0 and "Female" for 1.



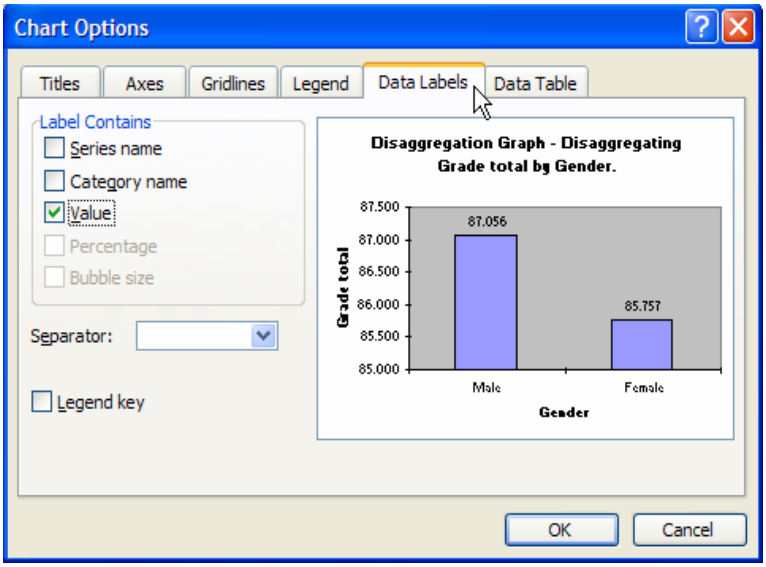
You should see that appropriate titles are now present for each bar!

STEP THREE



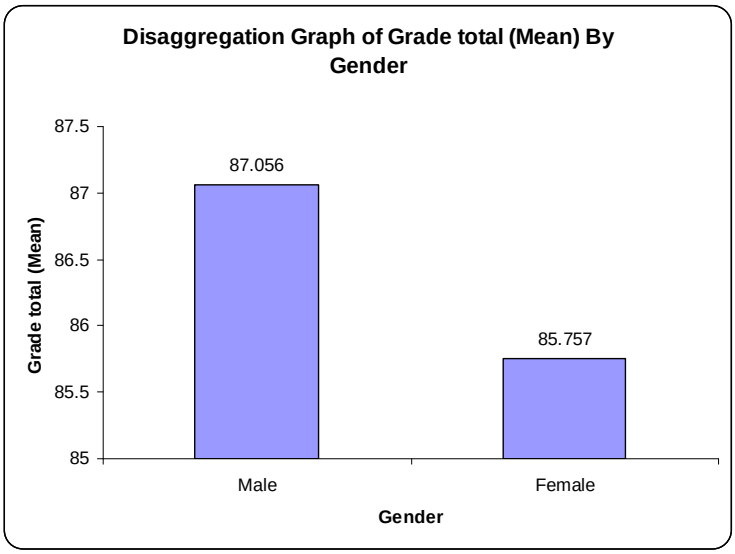
With the Disaggregation graph visible, select "Chart Options" from the menu bar

STEP FOUR



In the "Chart options" dialog box that appears, click on the "Data Labels" tab of the Chart Options dialog box and place a check mark next to "value"

EZAnalyze Tutorial 2



You should see an improved Disaggregation Graph!

NOTE – you can cut and paste this graph into Word or a PowerPoint Presentation!

FOR ADDITIONAL GRAPH MODIFICATIONS, SEE "MODIFYING GRAPHS" IN THE MANUAL

ANSWERING PAT'S QUESTION

Do male and female students have similar academic achievement?

To answer this question, we will use the EZAnalyze Disaggregation Graph to obtain the answer

Do male and female students have similar academic achievement?
This question can be answered by looking at the Bars of the Disaggregation graph.

ANSWER – Males and females have different academic achievement. The average grade for males is 87, while the average grade for females is 86.

WAIT A MINUTE! It *looks* like males and females perform very differently, but in reality, they are only about 1.4 points different according to our data. What is going on here?

This brings up an interesting point you should be aware of when interpreting graphs of this type. ALWAYS remember to think about what the *possible range of scores* is for the dependent variable. Here, the possible range of scores is 100 – in theory, scores can range from a zero to 100.

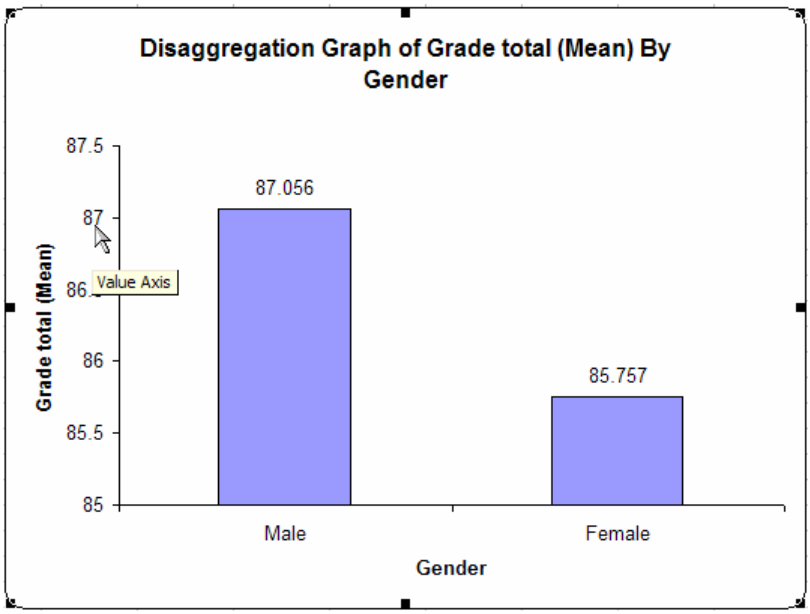
Extra Practice – you can perform an independent samples t-test to test the significance of the difference between males and females; go into the EZAnalyze *Advanced* menu, select *T-Tests,* then select *Independent samples t-test.*

EZAnalyze Tutorial 2

CREATING AN *HONEST* DISAGGREGATION GRAPH

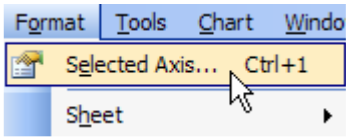
To make an *honest* graph, we need to change how the dependent variable is reported.

STEP ONE



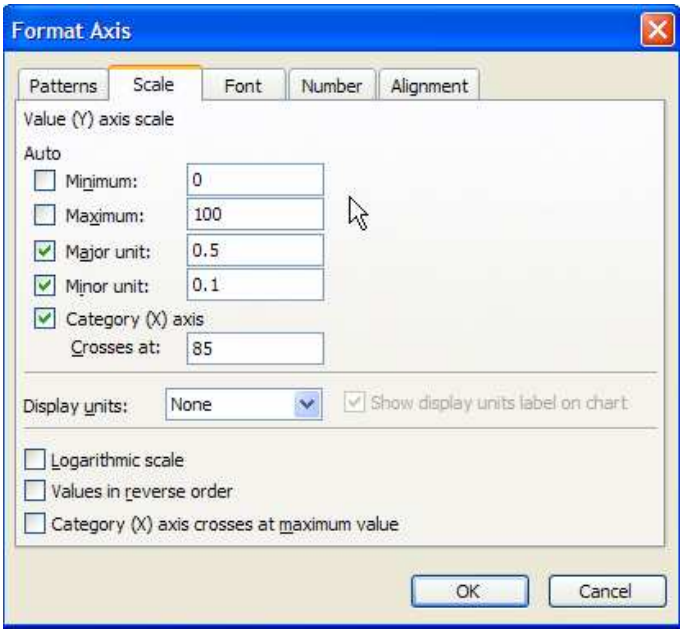
With the graph visible, click on the number "87" to select the "value axis"

STEP TWO



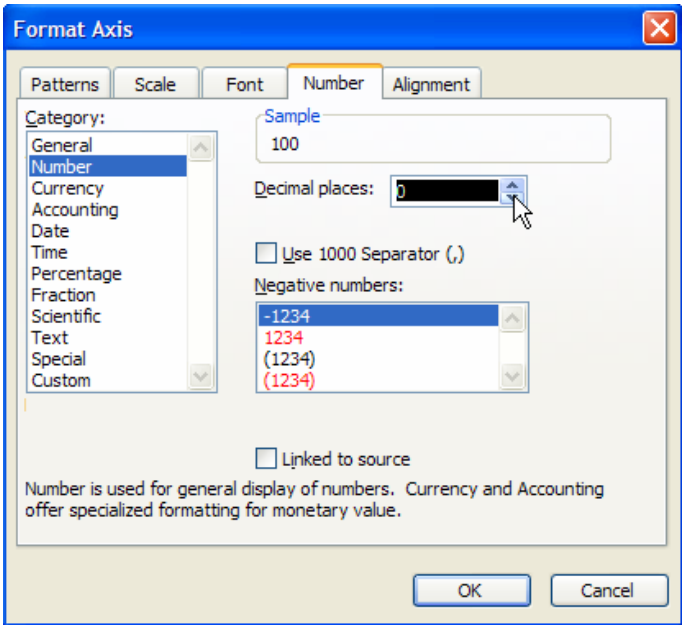
Select "Selected Axis" from the Format menu

STEP THREE

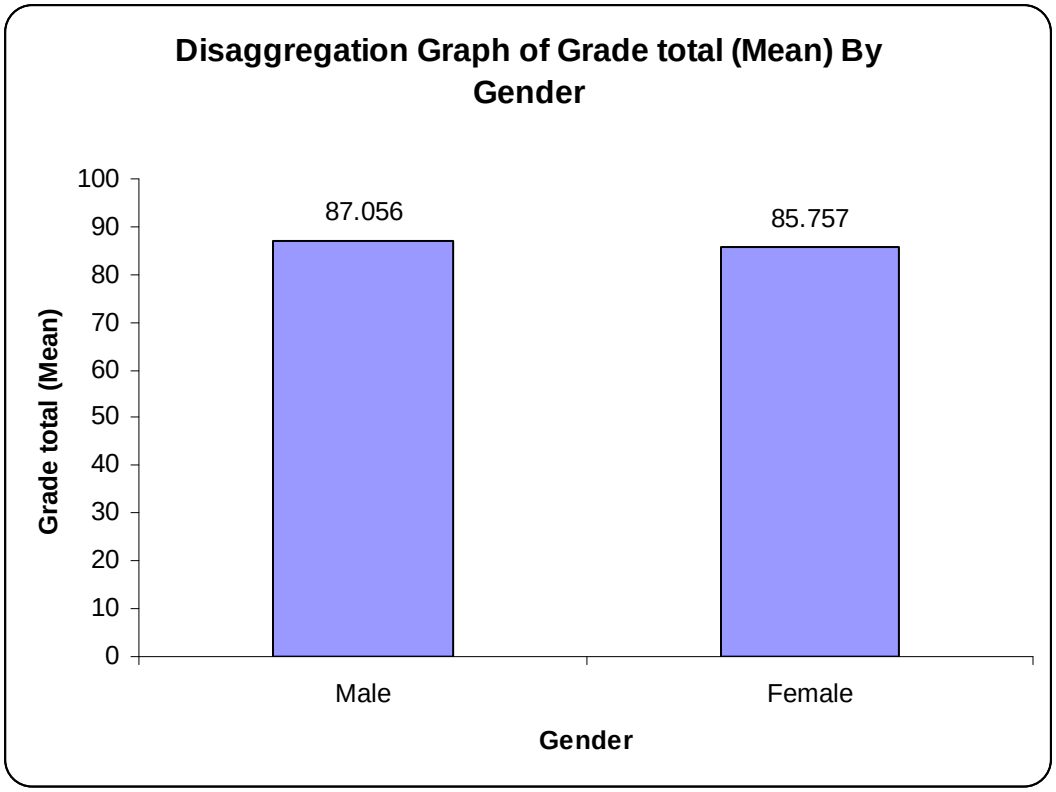


In the "Format Axis" dialog box that appears, change the "minimum" to "0" and the "maximum" to "100" – don't click OK yet!

STEP FOUR



Click on the "Number" tab, and change the number of decimal places from 3 to 0. Click OK when you are done



Now you have an honest *disaggregation graph!*

NOTE: You can now copy and paste this graph from Excel into a Word or PowerPoint document, or even into an HTML document!

The EZAnalyze Manual

The EZAnalyze Manual

SETTING UP YOUR DATA SHEET

For EZAnalyze to work smoothly, the Excel file that you use to provide the data for analysis needs to be set up with a few simple rules in mind. While there may be exceptions to these guidelines, following them will have you spending less time preparing your data, and more time analyzing it!

ROWS

Each row in your Excel data sheet (rows are numbered down the left side) should contain one "case." Usually, each row will contain data from one student or person. However, if your data is from schools in a district, and you are looking at how schools are different from one another, your "cases" would then be each individual school, and you would put each schools data into the rows. THE FIRST ROW IN YOUR DATA SHEET MUST CONTAIN VARIABLE NAMES!

COLUMNS

Each column in your Excel data sheet (columns are lettered across the top) should contain one "variable." EZAnalyze reads the variable names it will use for your analyses from the first row of data. The ability of EZAnalyze to read these variable names from the first row of data is what makes it special - and what makes it easy for you to analyze your data.

	A	B	C	D	E
1	name	grade	gender	english	math
2	juan valdez	6	1	99	98
3	julio iglesias	7	1	72	78
4	jane doe	8	2	90	82
5	john doe	6	1	88	45
6	john deere	7	1	76	55
7	johnny damor	8	1	55	66
8	jane deere	6	2	85	77
9	mr X	7	1	80	88
10	joe schmoe	8	1	67	66
11	george W	6	2	44	32

Sample of what your data should look like

Notice how the first row contains the variable names, and rows 2-11 contain the values for each case.

MISSING VALUES

Sometimes, you might have some data that is "missing." While you should make every effort to make sure that you have all of the data for each case, sometimes that is not possible. When this happens, you should leave the cell blank - do not put a number, letter, or anything in the cell where you have missing data. EZAnalyze will tell you how many missing values you have if you use the DESCRIBE - PERCENTAGES function. You can still calculate descriptive statistics using the DESCRIBE - DESCRIPTIVE STATISTICS function - the mean, median, and mode will be calculated using all of the data that are valid (not missing).

For specific information on how missing values affect the creation of new variables, see the manual topic associated with NEW VARIABLES.

SIZE

EZAnalyze can work with Excel data sheets that are 256 columns wide and over 65,000 rows long. If you are looking at data that contains more than 256 variables and 65000 cases, you should probably look into buying more powerful statistical analysis software such as SPSS.

NOTE: Even though Excel 2007 supports more than 256 columns and 65000 rows, EZAnalyze does not at this time.

DATA

EZAnalyze works best when all of the data (values contained in each cell) are numbers - HOWEVER, there are some exceptions. For example, if you have a variable named "gender" (see example above), you could represent "males" with a 1 and "females" with a 2. Since "gender" is what is known as a categorical variable (a variable that creates categories for naming purposes only – see the glossary), you will not be computing descriptive statistics such as means and standard deviations. In this case, then you can type "male" or "female" into your data sheet instead of numbers. TO BE CLEAR, however, you will always be safe if you use numbers.

The EZAnalyze Manual

DESCRIBING DATA

Using the Describe Function

To use the describe menu, select "Describe. . ." from the EZAnalyze main menu in Excel. You will be presented with two options – *Descriptive statistics* or *Percentages*.

DESCRIPTIVE STATISTICS

Descriptive statistics are commonly used statistics for reporting general characteristics of your data. Specifically, you can obtain the mean, median, mode, sum, range, minimum value, and maximum value for each numeric variable contained in your data.

Using the Describe - Descriptive Statistics Function

Descriptive statistics, which is part of the DESCRIBE menu, allow you to summarize your data in powerful yet easy to understand terms.

To obtain Descriptive statistics for your data, select the "Describe" option from the EZAnalyze menu in Excel, then select "Descriptive Statistics".

In the "Descriptive Statistics" dialogue box, select one or more *variables* from the variable list, which is located under the question "What would you like to describe?"

OPTIONS:

- Once you have selected the variables you would like to describe, you can select which statistics EZAnalyze will generate for you. The N is generated by default. You have the following options:

Mean	For each selected variable, reports the average for all of the cases in your sample
Median	For each selected variable, reports the value that falls halfway between the highest score and the lowest score
Minimum value	For each selected variable, reports what the lowest number is
Maximum value	For each selected variable, reports what the highest number is
Standard deviation	For each selected variable, reports how far on average each score deviates from the mean
Range	For each selected variable, reports what the total number of scores are (minimum value subtracted from maximum value)
Sum	The total of all scores in the variable added together

When you click OK, a results report will be printed on a separate sheet for your review.

For a walkthrough on this topic, see page 19

Reading your Results Report - Descriptive Statistics

Interpreting the results of the Describe - Descriptive Statistics function

The results report for your descriptive statistics contains information critical to the interpretation of your results. You have several options for describing your data (mean, standard deviation, range, etc); each of these are described above.

	A	B	C
1		**english**	**math**
2	**N Valid:**	10.000	10.000
3	**N Missing:**	0.000	0.000
4	**Mean:**	75.600	68.700
5	**Median:**	78.000	71.500
6	**Standard Deviation:**	16.754	20.205
7	**Range:**	55.000	66.000
8	**Minimum Value:**	44.000	32.000
9	**Maximum Value:**	99.000	98.000

Here is an example of a results report for two variables, "english" and "math."

The results report for this analysis is structured so that the columns contain all of the variables you selected for analysis, while the rows contain the descriptive statistics. In the sample, the first two rows of the results report "N valid" and "N missing" are always included in the results report; you are only required to have one of the remaining descriptive statistics included in your results report.

The EZAnalyze Manual

PERCENTAGES

Percentages allow you to describe your data in terms of the relative frequency of occurrence for each value of a given variable. For example, if you had a data set that contained a "gender" variable, calculating the percentages on the gender variable would tell you what percent of the people in your data set are males and what percent are females. ALSO, if you have missing values in your data, you can use the PERCENTAGES function to tell you how many missing values there are.

Using the Describe - Percentages Function

The DESCRIBE - PERCENTAGES function, part of the DESCRIBE MENU in EZAnalyze, allows you to summarize your data by both *frequency, percent, valid percent, and cumulative percent*.

To obtain percentages for your data, select the "Describe" option from the EZAnalyze menu in Excel, then select "Percentages".

In the "Percentages" dialogue box, select one or more *variables* from the variable list and click "OK"

When you click OK, a results report will be printed on a separate sheet for your review.

For a walkthrough on this topic, see page 16

Reading your Results Report – Percentages

Interpreting the results of the Describe - Percentages function

The results report for your percentages contains frequency and different kinds of percentage data for each variable you selected for analysis.

	A	B	C	D	E
1	***grade***				
2		**Frequency**	**Percent**	**Valid Percent**	**Cumulative Percent**
3	6	4.000	40.000	40	40
4	7	3.000	30.000	30	70
5	8	3.000	30.000	30	100
6					
7	***gender***				
8		**Frequency**	**Percent**	**Valid Percent**	**Cumulative Percent**
9	1	7.000	70.000	70	70
10	2	3.000	30.000	30	100

In this example, the variables "grade" and "gender" were selected for analysis.

Definitions for the columns are as follows:

Frequency. This is the total number of cases that there are for each level of the variable. In the example above, we can see that there are 4 students in 6th grade, and 3 each in 7th and 8th grade. If your variable had missing values, a row would be displayed indicating how many missing values there were.

Percent. This is the percent of the total number of cases that each level of the variable contains. In the example above, we can see that 70% of the students had a gender of 1 (male), while the remaining 30% had a gender of 2 (female). If your variable had missing values, a row would be displayed indicating the percent of the total number of cases that contained missing values.

Valid Percent. This column is only useful if you have missing values; otherwise, the results are identical to the "Percent" column. If you do have missing values, the Valid Percent column contains the percentage of the total for each level of the variable, not including missing values.

Cumulative Percent. This column simply adds up the valid percents to obtain a total of 100.

The EZAnalyze Manual

DISAGGREGATING DATA

The DISAGGREGATE function in EZAnalyze allows you to summarize your data by categories that are created by a categorical variable. Disaggregating data is a critical part of the No Child Left Behind legislation intended to close the "achievement gap," which is the differential academic performance of students according to their ethnic background. The achievement gap is an excellent example of disaggregated data, as it shows how differently students perform on standardized tests (the dependent variable) according to which ethnic group they belong (the categorical variable).

Using the Disaggregate function

To *disaggregate* your data, select the "Disaggregate." option from the EZAnalyze menu in Excel. "Double Disaggregation" is also explained below.

When the "Disaggregate" dialogue box appears, you will see two boxes with a list of all of the variables in your data set. The first box, labeled "Dependent variable(s)", is where you choose which variables you want summarized. You can choose one variable, or more than one. Some examples of common dependent variables in school settings are GPA, number of days absent, number of behavior referrals, and standardized test scores.

The second box, labeled "Categorical variable," is where you select the variable that tells EZAnalyze about your groups. This variable can be a *numeric variable* or a *string variable.* You can only select one categorical variable to disaggregate your data by from this list, although you may choose a second categorical variable to perform "Double Disaggregation", which is explained below. Some examples of common categorical variables in schools are ethnicity, gender, grade level, ELL status, socioeconomic status, and learning disability status.

OPTIONS:

- Once you have selected your dependent and categorical variables, you can then select descriptive statistic options located under the heading "statistic reported." The N is generated by default. Your options are:

Mean	The average or mean for the variable of the cases in your sample
Median	The score that falls halfway between the highest score and the lowest score for the variable
Standard deviation	How far on average each score deviates from the mean for the variable
Range	The total range of scores (minimum value subtracted from maximum value) for the variable
Minimum value	For each variable, reports what the lowest number is
Maximum value	For each variable, reports what the highest number is
Sum	The total of all scores in the variable added together

- **Double Disaggregation** - place a check mark in the box next to "Then by categorical variable 2" and select your second categorical variable from the drop-down box.

When you click OK, a results report will be printed on a separate sheet for your review.

Reading your Results Report - Disaggregate

Interpreting the results of the Disaggregate function

The results report for your disaggregation analysis contains several pieces of information that are important to properly interpret your results.

	A	B	C	D	E	F	G	H	I
1	*EZAnalyze Results Report - Grade total Disaggregated by Grade level*								
2									
3	Grade level		N	Range	Minimum	Maximum	Median	Mean	Std. Deviation
4	9	Grade total	41.000	37.130	62.150	99.280	89.870	87.742	7.832
5	10	Grade total	28.000	29.050	69.620	98.670	88.555	87.295	8.090
6	11	Grade total	25.000	30.230	69.310	99.540	87.650	87.440	7.353
7	12	Grade total	14.000	36.290	63.070	99.360	82.145	81.286	12.538

The results report is structured so that the very first row tells you what the analysis is - (dependent variable) Disaggregated by (categorical variable).

If you are double disaggregating your data, you will have two columns of categorical variables.

The remainder of the results report is a table - the first column (column A) in the table contains the levels of your categorical variable. In the example, the variable "Grade level" contained four levels - grade 9, grade 10, grade 11, and grade 12.

The second column contains the name of your dependent variable (Grade total in the example above).

The third column “N” contains the number of valid cases that there were for each level of the categorical variable.

The remaining columns contain the descriptive statistics you selected for the disaggregation analysis. Definitions are listed above and in the glossary.

For a walkthrough on this topic, see page 23

CREATING NEW VARIABLES

Using the New Variable Functions

The NEW VARIABLE functions in EZAnalyze allows you to easily create new variables from your existing data - you can create total scores or difference scores with the NEW VARIABLE function.

To create a new variable with your data, select the "New Variable" option from the EZAnalyze menu in Excel, then select the type of new variable you would like to create from the sub-menu. Each of the new variable types are described below.

Creating a New Summary Variable

Using the New Variable – Summary Function

The NEW VARIABLE - SUMMARY function, part of the NEW VARIABLE MENU in EZAnalyze, allows you to create a new variable that contains a new summary variable (either a *total score* or *average score*) quickly and easily - without having to calculate the new scores manually or by writing a formula to add to your Excel data sheet. An example of how you can summarize data with a new summary variable is for creating a total score for a survey you have created. If the survey has 10 items, you can create a new summary variable that contains either the average (mean) score or the total (all scores are added together) score for the survey.

To create a new summary variable for your data, select the "New Variable " option from the EZAnalyze menu in Excel, then select "Summary Variable."

In the "New Summary Variable" dialogue box, select more than one *variable* from the variable list. In the example provided above, you would select all 10 variables representing each question in your survey (hold down the CTRL key to select more than one variable) from the variable list.

Next, select a method for summarizing your data. If you select "Total," all of the variables you selected will be added together, and the result displayed as a new variable. If you select "Average," the average (or mean) score will be the result displayed as a new variable.

Next, give your new variable a name in the "Variable Name" box. This new name should be a variable name that has not already been used.

For a walkthrough on this topic, see page 9

Creating a New Difference Variable

Using the New Variable - Difference Function

The NEW VARIABLE - DIFFERENCE function allows you to create a new variable that contains a new difference score (glossary) quickly and easily - without having to calculate the new scores manually or by writing a formula to add to your Excel data sheet. Difference scores are very useful for showing changes over time with one number.

To create a new difference score variable for your data, select the "New Variable" option from the EZAnalyze menu in Excel, then select "Difference Variable."

In the "New Difference Variable" dialogue box, select one *variable* from the variable list under "Variable 1" and one variable from the variable list under "Variable 2." If you are creating a difference variable that is the difference between a pretest and a posttest, select the variable that represents the posttest from the "Variable 1" list, and the

variable that represents the pretest from the "Variable 2" list. The method that EZAnalyze uses for creating the difference score is to subtract "Variable 2" from "Variable 1." If your intervention was effective, subtracting pretest scores from posttest scores should give you positive numbers, which are a little easier to interpret. If you create a new variable and find that most of your numbers are negative, it might be a good idea to switch the variables around so that you are working with positive numbers.

Next, give your new variable a name in the "Variable Name" box. This new name should be a variable name that has not already been used.

When you click OK, a new column will be added to your data sheet that has the name you provided in the "Variable Name" box. The values contained in the cells of the new column will the difference score between the variables you have chosen.

For a walkthrough on this topic, see page 11

Creating a Percent Change Variable

Using the New Variable – Percent Change Function

The NEW VARIABLE - PERCENT CHANGE function, part of the NEW VARIABLE MENU in EZAnalyze, allows you to determine the difference between two variables in terms of the percent of change from baseline - for example, the amount of change that occurred between a pretest and posttest. This option is similar to the Difference Variable function, except that it provides a more standardized way of reporting the difference between the two variables.

To create a new percent change variable for your data, select "New Variable" from the EZAnalyze menu in Excel, then select 'Percent Change."

In the "New Percent Change Variable" dialogue box, select one *variable* from the variable list under "Baseline" and one variable from the variable list under "Second." If you are not sure which variable you should put first, try it both ways, and use the way that has more positive numbers than negative numbers as they are easier to interpret and work with.

Next, give your new variable a name in the "Variable Name" box. This new name should be a variable name that has not already been used.

OPTION:

- **Create more percent change variables** - check this box if you would like to create more of these variables. This is simply a time saver.

When you click OK, a new column will be added to your data sheet that has the name you provided in the "Variable Name" box. The values contained in the cells of the new column will the percent change for each individual score between the variables you have chosen.

Creating a New Standardized (Z) Score Variable

Using the New Variable – Standardized Score Function

The NEW VARIABLE - STANDARDIZED SCORE function, part of the NEW VARIABLE MENU in EZAnalyze, allows you to convert your data to 'standardized scores', or Z scores. Standardizing your variables is useful for putting things 'on the same metric'. For example, if you have two variables, and one is scored on a 5 point scale and the other is scored on a 7 point scale, they are difficult to compare side by side. If you standardize both variables, you can compare the standardized scores side by side easily. You can choose to create standardized scores based on a mean and standard deviation from your own data, or if the population parameters are known, you may choose to use those.

To create a standardized score variable for your data, select the "New Variable" option from the EZAnalyze menu in Excel, then select "Standardized (Z) Score."

In the "Standardize Score" dialogue box, select a *variable* from the variable list. In the example provided above, you would run this procedure twice - once for the variable on the 5 point scale, and the second time for the variable on the 7 point scale.

OPTIONS:

- **Use Mean and Standard Deviation from my data** - use this option if you want to standardize your scores based on the mean and standard deviation obtained from your sample's data
- **Use Mean and Standard Deviation I specify** - use this option if you have the 'population parameters' to standardize your sample's data on. Input those parameters into the appropriate spaces provided.
- **Create more standardized variables** - check this if you have more standardized variables to create. This is simply a time saver.

When you click OK, a new column will be added to your data sheet. This column will be named “Z_” plus the name of the variable you chose to standardize. This new variable will contain scores expressed in standard deviation units of the variable you selected in your original data.

Creating a New Percentile Rank Variable

Using the New Variable – Percentile Rank Function

The NEW VARIABLE - SUMMARY function, part of the NEW VARIABLE MENU in EZAnalyze, allows you to create a new variable that converts your data to their percentile rank equivalent. For example, if you want to know who is in the top 10% of the senior class at your high school, you can convert their overall GPA into a percentile rank variable to help you see who is in the 90th percentile or higher.

To create a new percentile rank variable (or variables) for your data, select the "New Variable" option from the EZAnalyze menu in Excel, then select "Percentile Rank."

In the "New Percentile Rank Variable" dialogue box, select more one or more *variables* from the variable list. In the example provided above, you would select the overall GPA from the variable list.

OPTION:

- **Number of decimal places** - use this to set how many decimal points are displayed. Setting this to zero does not 'round' the actual number to the nearest whole number; it simply rounds the number that is displayed. If you need more precise values than what is provided, you can format the cells in Excel to display as many decimal places as you would like.

When you click OK, a new column will be added to your data sheet. This column will be named “Rank_” plus the name of the variable you chose. This new variable will contain scores expressing the relative standing of scores on the variable you selected.

Creating a New Binary Variable

Using the New Variable – Binary Variable Function

The NEW VARIABLE - BINARY function, part of the NEW VARIABLE MENU in EZAnalyze, allows you to create a new variable (or variables) that are scored as either a 0 or a 1 - a process also known as 'dummy coding'. This is a very useful, and probably underutilized tool. For example, if you wanted to create a disaggregation graph using the percent of people who 'agreed' or 'strongly agreed' with your survey question, you can create a new binary variable that is scored a 1 if people selected 'agree' or 'strongly agree', and a 0 if they did not. This is useful because binary variables that are scored as a 0 or a 1 have a 'special property', and that special property is that the mean of a binary variable with values of 0 or 1 is the percent of people who scored a one.

To create a new binary variable for your data, select the "New Variable" option from the EZAnalyze menu in Excel, then select "Binary Variable".

In the "New Binary Variables" dialogue box, select one *variable* to convert from list on the left. When you click on a variable, you will see the list under 'Current Values' populate. Each value you select will be converted into a new binary variable on your data sheet.

OPTIONS:

- **Create a single binary variable...** - check this box if you would like all of your selected values to be converted into a single binary variable. In the example above, you would select values 4 and 5, and check this box to create the appropriate binary variable.
- **Create more binary variables** - check this box if you are planning on making more binary variables. It is simply a time saver.

When you click OK, a new column will be added to your data sheet for each value you selected in the 'current values' box. Each new variable will have the original variable name, '_Bin_' and the selected value as the new variable name.

Creating a New Random Numbers Variable

Using the New Variable – Random Numbers Function

The NEW VARIABLE - RANDOM NUMBERS function, part of the NEW VARIABLE MENU in EZAnalyze, allows you to quickly create random numbers to demonstrate various statistical problems and concepts - or 'just for fun'. You can set the mean and standard deviation, generate completely random numbers within a specified range, or specify a range, mean, and standard deviation.

To create a random numbers variable for your data, select the "New Variable" option from the EZAnalyze menu in Excel, then select "Random Numbers."

In the "Random Number Generator" dialogue box, type a name into the 'Variable name' box; this is what will appear in the first row on your data sheet. Also, specify the number of rows of data you would like to generate.

Next, choose at least one of the first two options below

OPTIONS:

- **Specify the range of the random numbers** - use this option if you want to restrict the random numbers to a certain range by specifying the minimum and maximum values you would like to allow
- **Specify the mean and SD of the Random Numbers** - use this option to see how randomly generated scores cluster around a given mean and standard deviation. This will produce a relatively normal distribution if the amount of random numbers you chose to generate is reasonable. But, they are random after all...
    - Whole numbers only - check this if you dont want decimals in your variable
- **Create more random variables** - check this if you have more random number variables to create. This is simply a time saver.

When you click OK, a new column will be added to your data sheet that has the name you specified, with randomly generated numbers with parameters you specified.

CREATING GRAPHS

Using the Graph Function

The GRAPH function in EZAnalyze allows you to create visual representations of your data with three common types of graphs - Pie Charts, Bar Charts, *and* Area Charts*. These types of graphs can provide you with powerful pictures of your data, and provide you with a meaningful way to communicate with other people about the results of your data analyses.*

To create a graph with your data, select the "Graph" option from the EZAnalyze menu in Excel. The sub-menu presents you with three graphing options:

Histogram. A histogram is a graph that shows you the *distribution of scores* for a variable. Histograms can be used for all types of numeric and categorical variables, and show you how many cases (as a number, or percent of the total number of cases) there are within a certain range of scores. For example, if you wanted to know how many students there were in each ethnic group in your school, you could use a histogram to create a graph that would show you how many students there were in each group. You have several choices for creating histograms. A useful choice to demonstrate statistical concepts is the 'Traditional Histogram' with multiple variables, which allows you to display the distribution for several variables on one graph.

Disaggregation. A disaggregation graph allows you to create a graph of disaggregated data. These graphs can be particularly powerful to show differences between groups.

Multiple Variable. A multiple variable graph allows you to display the mean, sum, or N for several variables side-by-side. Here is where you can also create 'Error Bar Graphs', which are useful for demonstrating statistical concepts.

Creating Histogram Graphs

Using the Graph - Histogram Function

The GRAPH - HISTOGRAM function, part of the GRAPH MENU in EZAnalyze, allows you to summarize your data by showing how people scored on a given variable. This could be something as simple as a pie graph depicting the percentage *of students of each ethnic background or a set of histograms overlaid ontop of each other.*

To obtain a histogram for your data, select the "Graph" option from the EZAnalyze menu in Excel, then select "Histogram".

In the "Histogram" dialogue box, select the type of histogram you would like to create:

Pie Chart - this will create a pie chart of a single variable you select.
Bar Chart - this will create a bar chart of a single variable you select.
Area Chart - this will create an area chart of a single variable you select
Traditional Histogram - this will create a graph with a 'smoothed line' of one or more variables you select.

OPTION:

- **Advanced Option** - you can select the number of categories you would like displayed along the X Axis (bottom) of the chart. If there fewer than 10 categories in your selected variable(s), all categories will be displayed. If there are more than 10 categories (for example, if you are graphing GPA or number of days absent), it may be useful to change this setting to make your graph more interpretable.

When you click OK, EZAnalyze will create a new sheet - a results report containing the graph, and the data used to make that graph.

For a walkthrough on this topic, see page 27

Reading your Results Report - Histogram Graphs

Interpreting the results of the Graph - Histogram function

Now that you have created your histogram, it is time to interpret what it is telling you. When you created your histogram, a "chart" (your graph) was created, in addition to the data used to make the chart. These two pieces work together; if you change your data, the chart will also change.

You are provided with three options for creating a histogram - a pie chart, a bar chart, an area chart, or a 'Traditional Histogram'.

Pie Chart. A pie chart is relatively easy to interpret. The entire pie represents all possible values (except *missing values*) contained in the variable that you selected for the histogram analysis. Each slice of the pie represents the percentage of the total that each category contained in your data. In the "Legend" located on the right side of the pie chart, you will see what each slice of the pie represents. The only "problem" with pie charts is that they do not tell you how many people there were in each category (although you can show this in the Chart Options in Excel).

Bar Chart. A bar chart is also relatively easy to interpret. The bottom of the chart (called the *X Axis*) contains the categories or all possible values (except *missing values*) in the variable you selected for histogram analysis. The left side of the chart (called the *Y Axis*) contains the *frequency* of occurance - a simple count of how many times the specified value occurred. If your data sheet contains individual student data, the Y Axis will tell you how many students there were in each category.

Area Chart. An area chart is really nothing more than a fancy bar chart - it is interpreted precisely the same way.

Traditional Histogram. This type of graph can be interesting, and somewhat more complicated to interpret, because you can include more than one variable on the histogram. At its core, this type of graph is called 'traditional' because of the smooth line that connects the values along the X axis. You will see the various values of your chosen variable along the bottom (X axis), and the frequency count along the left side (Y axis).

To make your charts look prettier, please see the help topic associated with *Modifying your graphs*.

For a walkthrough on this topic, see page 27

Creating Disaggregation Graphs

Using the Graph - Disaggregate Function

The GRAPH - DISAGGREGATE function, part of the GRAPH MENU in EZAnalyze, allows you to display disaggregated data visually. If you would like to learn more about disaggregation, start with learning how to use the DISAGGREGATE function of EZAnalyze. Examples of useful ways to use a disaggregation graph would be to show how boys and girls score on a standardized test, to depict how different ethnic groups have different dropout rates, or to show how dropout rates differ by ethnicity and gender using "double disaggregation."

To obtain a disaggregation graph for your data, select the "Graph" option from the EZAnalyze menu in Excel, then select "Disaggregation".

In the "Disaggregation Graph" dialogue box, you will see two variable lists. In the variable list located under the heading of "Dependent variable" select one or more variables from the list (NOTE: you can only choose one if you want to double disaggregate). In the next variable list box located under the heading "Categorical variable," choose the *categorical variable* that you would like to disaggregate your *dependent variable*(s) by. Typical categorical variables in education are ethnicity, gender, ELL status, and grade level.

OPTIONS:

- **Statistic Reported** - Choose the mean to obtain an average score, the sum to obtain the total summed score, or the N for each group in your categorical variable
- **Graph Type** - Choose a Bar, Area, or Line graph
- **Double Disaggregation** - Place a check in the box next to 'Then by categorical variable 2', and then select a second categorical variable. NOTE: You can not double disaggregate multiple dependent variables)

To create a "Double Disaggregation" graph of your data, place a check mark in the box next to "Then by categorical variable 2" and select your second categorical variable from the drop-down box.

Next, select the "Statistic Reported." This is how your dependent variable will be summarized. Selecting the mean will give you the average score on the dependent variable for each group, selecting the sum will give you the total summed score for each group, and selecting the N will tell you how many people there were in each group.

Next, select the type of graph you would like to create and click "OK". Bar charts and area charts are the most popular types of graphs for displaying disaggregated data; you might want to create both types of charts and choose the one you think is most appealing to you. If you are double disaggregating your data and select "area/line chart", a line chart will be displayed. This is an excellent way to display double disaggregated data.

When you click OK, EZAnalyze will create a new sheet - a results report containing the graph, and the data used to make that graph.

For a walkthrough on this topic, see page 32

Reading your Results Report - Disaggregation Graphs

Interpreting the results of the Graph - Disaggregation function

Now that you have created your disaggregation graph, it is time to interpret what it is telling you. When you created your graph, two new sheets were added to Excel - one sheet containing the chart, and another sheet containing the data for the chart. These two sheets work together; if you change data in the chart data sheet, the chart will also change. For example, if you change the categories (in the first column) from the numbers 1 and 2 to "male" and "female," you will have the new names defining your categories instead of numbers. Many times, this is much better!

You are provided with two options for creating a disaggregation graph - a bar chart, or an area/line chart. Bar charts are usually the best - but the others can also be useful. If you want to show changes over time (if your categories are different points in time - for example, semesters and years of high school) an area chart is very good. If you double disaggregate your data, you can use a "clustered bar chart" by selecting the bar chart option, or can use a line chart by selecting the area/line chart option. Graphs are often a matter of taste - play with it until you find one you like!

Bar Chart. A bar chart is relatively easy to interpret. The bottom of the chart (called the X Axis) contains the values of the categorical variable you selected for your disaggregation analysis. The left side of the chart (called the Y Axis) contains the Average or Mean score of the dependent variable you selected for disaggregation analysis. (Glossary)

Area Chart. An area chart is really nothing more than a fancy bar chart - it is interpreted precisely the same way.

Line Chart. A line chart is ONLY created for double disaggregated data. A separate line is created for each level of your second categorical variable. This can be VERY USEFUL for displaying double disaggregated data.

To make your charts look prettier, please see the help topic associated with *Modifying your graphs.*

For a walkthrough on this topic, see page 31

Creating Multiple Variable Graphs

Using the Graph – Multiple Variable Function

The GRAPH – MULTIPLE VARIABLE function, part of the GRAPH MENU in EZAnalyze, allows you to compare several different variables side-by-side. You can create line, bar, or error bar graphs to compare your selected variables. Error bar graphs are particularly interesting, as they provide a picture of not only the average score, but the "spread" of the scores as well, using the standard deviations of the selected variables.

To obtain a Multiple Variable graph for your data, select the "Graph" option from the EZAnalyze menu in Excel, then select "Multiple Variable".

In the "Multiple Variable" dialogue box, select your variables from the variable list.

OPTIONS:

- **Choose the type of graph you would like to create** - choose one of the 4 available options
  - Graph of Means - select this option to graph the average of your selected variables
  - Graph of Sums - select this option to obtain the summed total for each of your selected variables
  - Graph of N's - select this option to graph the count of valid scores in your selected variables
  - Error Bar Graph - select this option to create Error Bar Graphs of your selected variables
    - If you select Error Bar Graph, you will be given the option to set the number of standard deviations above and below the mean.
- **Choose the type of graph** - Select bar or line graph according to your preference

---

Reading your Results Report – Multiple Variable Graphs

Interpreting the results of the Graph – Multiple Variable function

The Multiple Variable graph you created allows you to compare several variables side by side. Interpreting bar and line graphs are similar, while interpreting error bar graphs is just a little different.

Bar/Line Graphs. Interpreting your bar or line graph is relatively straightforward; along the X axis (bottom) of your graph are the variables you selected, while the Y (side) is used to display the scores (mean or sum) for each selected variable. It is important that these variables be on the same metric to ensure you are “comparing apples to apples.”

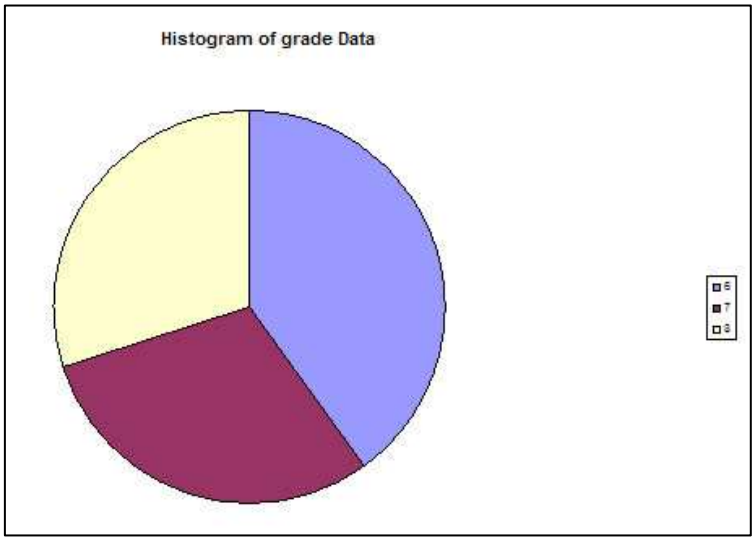
Error Bar Graph. Error bar graphs present your results a little differently; the X axis still contains your variables, and the Y axis contains the mean score. You can find the mean score by locating the box (usually near the middle of the line). Each “error bar” extends above and below the mean; in this case, the error bars extend the specified number of standard deviations above and below the mean score for each variable. This allows you to see not only the mean score for each variable, but how “spread out” scores on each variable are.

MODIFYING YOUR GRAPHS

General Graph Modification Guidelines

EZAnalyze includes some simple graphing functions to help get you started creating your own graphs. For detailed information on modifying how your graphs look, you should turn to the help documentation that is included in Excel. You can find this information by going to the HELP menu of Excel (Should be located next to the EZAnalyze menu) and looking in the Table of Contents for "charts and graphics."

To get you started, however, here is a brief overview of how you can make your charts and graphs look better! First, you should know that the "Chart" and "Data" in the sheet containing your results report, created when you clicked OK, work together. To illustrate how you can make changes to your chart, we will use the following example of a pie chart of a variable named "grade" that contains three categories - sixth grade, seventh grade, and eighth grade. The same principles apply for modifying bar charts and area charts.



Example of a histogram, unmodified

As you can see from the example, our histogram is not very pretty. Let's do some work to make it more appealing!

Changing your legend. You can change the legend by altering the data that was created with the chart. For the above example, it looks like this:

	A	B	C
1	6	4	
2	7	3	
3	8	3	
4			

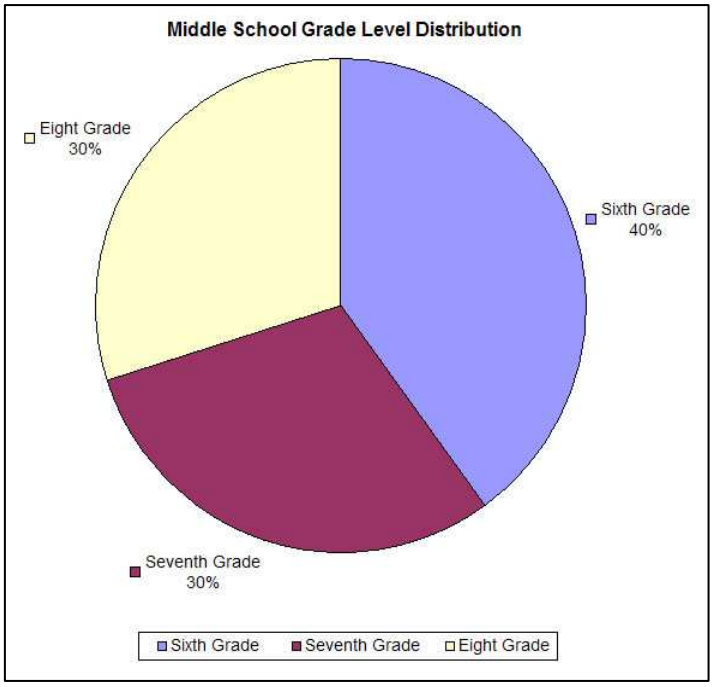
Column A contains our grade levels - 6, 7, and 8, and column B contains the number of students in each grade. We can replace the numbers in column A with words to have our legend make more sense.

Next, we will return back to the "Chart" that contains our graph by clicking on it. It is difficult to see what the percent of students in each grade is without numbers.

To add numbers, go to the CHART menu near the top of your screen and click on OPTIONS. In the "Chart Options" dialogue box, click on the tab that says "Data Labels" and place check marks in the boxes next to "Category Name", "Percentage", and "Legend Key". DO NOT CLICK OK YET!

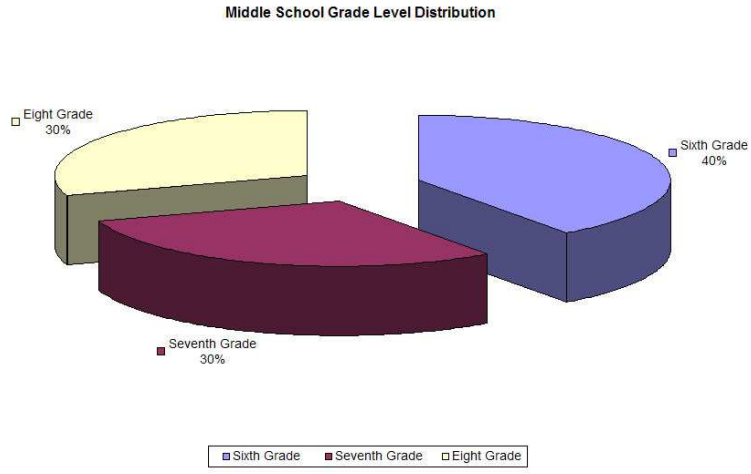
Next, we should give our pie graph a better name. In the "Chart Options" dialogue box, click on the tab that says "Titles." In the "Chart Title" box, erase the text that is there and give our graph a better name - here, we will type in "Middle School Grade Level Distribution." DO NOT CLICK OK YET!

Next, click on the "Legend" tab in the chart options box and choose a different place to put the legend. Click OK!



Voila! Our new, much prettier graph!

Want to add a 3-D effect? Go back into the CHART menu and select CHART TYPE. You will be given several 3-D options. Below is the "Exploded pie with 3-D visual effect" option.



Bar and Area Graph titles. EZAnalyze uses your variable names as the X and Y axis titles by default. Often, you will want to change these! To set meaningful titles on your bar graph, click once anywhere on your graph, go into the CHART menu and select CHART OPTIONS. On the "Titles" tab, type a new title for your categorical variable in the "Category (X) Axis" box and a new title for your dependent variable in the "Value (Y) Axis" box if you like.

Your are encouraged to "play around" with these chart options. If you mess something up, you can always go back to your original data sheet and create a new graph from scratch. You can also get to many of the graphing options by simply double-clicking on the area you would like to modify - give it a try!

For walkthroughs on this topic, see page 36

The EZAnalyze Manual

EZANALYZE ADVANCED FUNCTIONS

This section briefly covers how to perform the advanced functions, and how to read the results reports. They are not covered with as much depth as the other sections of the manual.

Correlation

Using the Correlation Function

The CORRELATE function in EZAnalyze allows you to see how two or more variables are related to each other. A correlation coefficient*, simply put, is a number between -1 and +1 that describes the direction and degree of relationship between two variables. The direction of the relationship is indicated by the sign (positive or negative, + or -), while the degree of the relationship is indicated by the number itself, which is usually a decimal between 0 and 1. The higher the number, the stronger the relationship is between two variables - 0 would indicate that there is no relationship, while 1 would indicate that there is a perfect relationship. For example, if you obtained a correlation coefficient of .41 between the number of years of schooling and gross income, that would indicate that there was a positive relationship between the number years of schooling and gross income. Alternatively, if the correlation coefficient was -.41, that would indicate that there was a negative relationship. For more information on interpreting correlation coefficients, see the associated topic for interpreting CORRELATE results reports.*

To obtain correlation coefficients with your data, select "Advanced" from the EZAnalyze menu in Excel, and then choose "Correlation".

When the "Correlate" dialogue box appears, you will be presented with two lists of the variables in your data sheet.

OPTIONS:

- **Correlate two variables and create a scatterplot** - Choose this option if you would like to obtain a single correlation coefficient which displays a graph of the results in addition to the coefficient and related statistics. Select the variables you would like to correlate from the 'Variable 1' and 'Variable 2' lists.
    - Show trendline on scatterplot - select this if you would like a 'line of best fit' displayed on the graph
- **Correlate more than two variables to create a correlation matrix** - Choose this option if you want to conduct all possible correlations between two or more variables. Select your variables from the list provided.

A NOTE ON SELECTING VARIABLES. While EZAnalyze will let you choose any two numeric variables to correlate, variables that are categorical, such as gender, ethnicity, or learning disability status, are not appropriate to conduct correlation analyses with in EZAnalyze. Examples of common correlations of interest to educators are the number of days absent with GPA, standardized test scores and GPA, and the number of disciplinary referrals with standardized test scores.

For a walkthrough on this topic, see page 25

Reading your Results Report – Correlations

Interpreting the results of the Correlation Function

The results report for your correlation contains a few pieces of information that are critical to interpreting your results report for the Correlate function. There are two different kinds of results reports, depending on whether or not you chose to create a correlation matrix of multiple variables, or a a single correlation coefficient with a scatterplot graph.

Results Report for Correlating Two Variables with a Scatterplot

First, the top row of your results report tells you which two variables were used in the correlation analysis - it is reported as ***Correlation of (Variable name 1) and (Variable name 2).***

You will also see three statistics reported - the Pearson Correlation coefficient, the N (number of valid pairs in your data), and the P values (significance level reached).

Next, take a look at the table that EZAnalyze has created - you will see that the variable names that you correlated are along both the top of the table (row 3) and along the side of the table (column B). To find the *correlation*

coefficient (Glossary), find the number located at the intersection of row 7 and column D next to the words *Pearson Correlation.* The *N* in the results report indicates how many cases were used in computing the correlation, and the P is the significance level.

You will also see a scatterplot showing the relationship between your two variables, and a trendline (line of best fit) if you selected that option.

Correlation Matrix Results Report

The correlation matrix results report provides the Pearson Correlation coefficient, number of valid pairs of data (N), and the significance level (P) of all possible correlations between the variables you selected.

RULES OF THUMB FOR INTERPRETING CORRELATION COEFFICIENTS.

1. The larger the correlation coefficient, the better.

2. A correlation of less than .3 is typically considered "not meaningful," a correlation between .3 and .5 is typically considered "weak," a correlation between .5 and .7 is typically considered "moderate," and a correlation that is greater than .7 is typically considered to be a strong correlation.

3. CORRELATION DOES NOT IMPLY CAUSATION. Because two variables are correlated, even if they are statistically significant, does not mean that one variable causes changes in the other - it means nothing more than "they are related."

T-TESTS

Choosing a T-Test Function

EZAnalyze contains three t-test options; a one sample, paired samples, or independent samples t-test. Each is briefly described below.

One Sample T-Test - appropriate for use when you are comparing a value that is known in the population to your data. For example, if you want to know how students in your school compare to students in the US on the SAT, you would use this test because we know the national average SAT score is 1000 (or perhaps 1500 if you use all three sections).

Paired Samples T-Test - use this when you want to compare data that are in two columns from the same person. This is a "repeated measures" t-test because appropriate use requires that the same person contributed scores to both measures. An excellent example of this is to see if there is a difference between pretest scores and posttest scores after some kind of intervention.

Independent Samples T-Test - the most common type of t-test, this is appropriate for use when you want to see how two different groups performed on some dependent measure. For example, if you want to see if boys and girls have significantly different attendance rates at your school, you would use the independent samples t-test to get your answer.

One Sample T-Test

Using the One Sample T-Test Function

The one sample t-test in EZAnalyze allows you to determine the statistical significance of a difference between an observed group mean in your data, and the value of a mean known to be true in a population. Since this is an Advanced feature of EZAnalyze, it will not be discussed at length here. Feel free to consult a statistics textbook or the Internet for information on what a one samplet-test is and how it works. If you are already familiar with one sample t-tests and want EZAnalyze to crunch the numbers for you, read on!

To conduct this analysis, your data will need to have only one variable. You will also need to have a "Numerical Test Value" (NTV) to tell EZAnalyze what to compare your data to.

To conduct a one sample t-test, select "Advanced" from the EZAnalyze menu in Excel, then choose "T-Tests" and select "One sample".

In the "One Sample T-Test" dialog box, select a "Test variable" from the list of variables in your data set, and put a Numerical Test Value to compare your data to in the space provided.

When you click OK, a results report will be printed on a separate sheet for your review.

Reading your Results Report - One Sample T-Test

Interpreting the results of the One Sample T-Test function

The results report for your One Sample T-Test contains several pieces of information to help you interpret your results. These are not discussed at length here, but you are encouraged to consult a statistics textbook or the Internet to learn more!

The EZAnalyze results report for the one sample t-test contains descriptive statistics for your test variable, telling you what the mean, standard deviation, and number of cases in the analysis were. You are also provided with the mean difference (Mean Difference), a t-score, Eta Squared, and the significance of the difference between means (P). "NTV" is the Numeric Test Value you selected for the analysis.

Row 14 contains a statement indicating "in English" whether or not the observed difference between the NTV and your test variable were significant.

NOTE: If you are using a known group mean as the NTV, you should also know what the standard deviation is for that group mean. A one sample t-test estimates the standard deviation from your data, so check to make sure the standard deviation in your sample is similar to the known standard deviation in the population!

Further discussion of what these are is beyond the scope of this manual!

EZAnalyze also automatically creates a graph of your results. To learn more about modifying this type of graph, please see *Modifying your graph* for more information.

Paired Samples T-Test

Using the Paired Samples T-Test Function

The paired samples t-test in EZAnalyze allows you to determine the statistical significance of a difference between two paired means. This test is the most appropriate when you have a repeated measures *design where the data are from a single group of cases exposed to two measures over time, as in a pretest-posttest design. Since this is an Advanced feature of EZAnalyze, it will not be discussed at length here. Feel free to consult a statistics textbook or the Internet for information on what a paired t-test is and how it works. If you are already familiar with paired samples t-tests and want EZAnalyze to crunch the numbers for you, read on!*

To conduct this analysis, your data will need to have at least two "paired" variables - each variable should have similar data. Think of each of these as a dependent variable. What this test does is tell you if the mean of the first variable is significantly different from mean of the second variable.

To conduct an independent samples t-test, select "Advanced" from the EZAnalyze menu in Excel, then choose "T-Tests" and "Paired Samples".

In the "Paired Samples T-Test" dialog box, select your paired variables - one variable under "Variable 1" and another under "Variable 2" and click OK. If you are running this test on a pre-post test experimental design, you will want to use Variable 1 for your pretest variable, and Variable 2 for your posttest variable.

A note about your variables. These variables should be on the same metric - for example, GPA for social studies and science, or the mean scores of a variable coded from 1-5. This procedure will not be accurate if your variables are not on the same metric - for example, comparing GPA and SAT scores is not appropriate in a paired samples t-test.

When you click OK, a results report will be printed on a separate sheet for your review.

Reading your Results Report - Paired Samples T-Test

Interpreting the results of the Paired T-Test function

The results report for your Paired Samples T-Test contains several pieces of information to help you interpret your results. These are not discussed at length here, but you are encouraged to consult a statistics textbook or the Internet to learn more!

The EZAnalyze results report for the paired samples t-test contains descriptive statistics for each variable, telling you what the mean, standard deviation, and number of cases for your analysis were. You are also provided with the mean difference (Mean of Diffs), the standard error of the difference (SE of diff), Eta Squared, a t-score, and the significance of the difference between means (P).

Row 14 contains a statement indicating "in English" whether or not the observed difference between your paired variables is statistically significant.

Further discussion of what these are is beyond the scope of this manual!

EZAnalyze also automatically creates a graph of your results. To learn more about modifying this type of graph, please see *Modifying Your Graphs* in the manual.

Independent Samples T-Test

Using the Independent Samples T-Test Function

The independent samples t-test in EZAnalyze allows you to determine the statistical significance of a difference between two group means. Since this is an Advanced feature of EZAnalyze, it will not be discussed at length here. Feel free to consult a statistics textbook or the Internet for information on what an independent t-test is and how it works. If you are already familiar with independent samples t-tests and want EZAnalyze to crunch the numbers for you, read on!

To conduct this analysis, your data will need to have at least two variables - one dependent variable, and one categorical variable containing the groups you would like to test. NOTE - your categorical variable can have more than two groups - you select which specific groups you want to look at later.

To conduct an independent samples t-test, select "Advanced" from the EZAnalyze menu in Excel, then choose "T-Tests" and select "Independent samples".

In the "Independent Samples T-Test" dialog box, select a dependent variable in the left column under "Dependent variable" and an independent (grouping) variable in the middle column under "Categorical variable."

Once you have selected a categorical variable, you will see a list of all of the values contained in the categorical variable you selected. If you have two groups in your data, they are selected for you. Otherwise, select two of these values by clicking on each one once.

When you click OK, a results report will be printed on a separate sheet for your review.

Reading your Results Report - Independent Samples T-Test

Interpreting the results of the Independent Samples T-Test function

The results report for your Independent Samples T-Test contains several pieces of information to help you interpret your results. These are not discussed at length here, but you are encouraged to consult a statistics textbook or the Internet to learn more!

The EZAnalyze results report for the independent samples t-test contains descriptive statistics for each group, telling you what the mean, standard deviation, and number of cases in each group are. You are also provided with the mean difference, a t-score, Eta Squared, and the significance of the difference between means (P).

Row 13 contains a statement indicating "in English" whether or not the observed difference between group means is statistically significant.

Further discussion of what these are is beyond the scope of this manual!

EZAnalyze also automatically creates a disaggregation graph of your results. To learn more about modifying this type of graph, please see *Modifying Your Graphs* in this manual.

ANOVA – Single Factor

Using the ANOVA – Single Factor Function

The ANOVA function in EZAnalyze allows you to conduct a statistical hypothesis test on more than two group means. Since this is an Advanced feature of EZAnalyze, it will not be discussed at length here. Feel free to consult a statistics textbook or the Internet for information on what ANOVA is and how it works. If you are already familiar with ANOVA and want EZAnalyze to crunch the numbers for you, read on!

To ensure that your data is ready for an ANOVA, make sure that you have a single dependent variable in mind and that your categorical variable contains all of the groups. SO, for example, if you wanted to see if different ethnic groups in your school had significantly different GPA's, you would have one variable in your data to represent GPA, and another variable to represent Ethnicity.

To perform an ANOVA on your data, select "Advanced" from the EZAnalyze menu in Excel, and then choose "ANOVA".

When the "ANOVA" dialog box appears, select a dependent variable from the list on the left, and a categorical (independent) variable (your grouping variable) from the list on the right. When you are satisfied with your choices, click OK.

When you click OK, a results report will be printed on a separate sheet for your review.

Reading your Results Report – Single Factor ANOVA

Interpreting the results of the ANOVA – Single Factor Function

The results report for your ANOVA contains several pieces of information to help you interpret your results. These are not discussed at length here, but you are encouraged to consult a statistics textbook or the Internet to learn more!

The EZAnalyze ANOVA report contains descriptive statistics regarding the overall mean (the Grand Mean), the means and standard deviations of your subgroups, an ANOVA table listing sources of variance and the degrees of freedom, Eta Squared, and F and P values.

At the end of your results (before the graph), there is a statement "in English" telling you whether or not any significant differences were observed between any of the groups.

Simple Post Hoc tests are also conducted if there are more than two groups, and the P value for the ANOVA is less than .05. The post hoc tests provide the mean difference, T score, unadjusted and Bonferroni adjusted P values, and Eta Squared for all possible pairwise comparisons.

EZAnalyze also automatically creates a graph of your results. To learn more about modifying this type of graph, please see *Modifying Your Graphs*.

Repeated Measures ANOVA

Using the Repeated Measures ANOVA Function

The Repeated Measures ANOVA function in EZAnalyze allows you to conduct a statistical hypothesis test on more than two within-subject means. Since this is an Advanced feature of EZAnalyze, it will not be discussed at length here. Feel free to consult a statistics textbook or the Internet for information on what a Repeated Measures ANOVA is and how it works. If you are already familiar with Repeated Measures ANOVA and want EZAnalyze to crunch the numbers for you, read on!

To ensure that your data are ready for a Repeated Measures ANOVA, make sure that you have at least two variables that are on the same metric and contain “within-subjects” data.

To perform a Repeated Measures ANOVA on your data, select "Advanced" from the EZAnalyze menu in Excel, and then choose "Repeated Measures ANOVA".

When the "Repeated Measures ANOVA" dialog box appears, select your variables for analysis from the list.

You can also choose the type of graph (line or bar) you would like EZAnalyze to create. When you are satisfied with your choices, click OK.

When you click OK, a results report will be printed on a separate sheet for your review.

Reading your Results Report – Repeated Measures ANOVA

Interpreting the results of the Repeated Measures ANOVA function

The results report for your Repeated Measures ANOVA contains several pieces of information to help you interpret your results. These are not discussed at length here, but you are encouraged to consult a statistics textbook or the Internet to learn more!

The EZAnalyze Repeated Measures ANOVA report contains descriptive statistics regarding the the means and standard deviations of your selected variables, an ANOVA table listing sources of variance and the degrees of freedom, Eta Squared, and F and P values.

At the end of your results (before the graph), there is a statement "in English" telling you whether or not any significant differences were observed between any of the groups.

Simple Post Hoc tests are also conducted if there are more than two Repeated Measures, and the P value for the ANOVA is less than .05. The post hoc tests provide the mean difference, T score, unadjusted and Bonferroni adjusted P values, and Eta Squared for all possible pairwise comparisons.

EZAnalyze also automatically creates a graph of your results. To learn more about modifying this type of graph, please see *Modifying Your Graphs*.

Chi Square

Using the Chi Square Function

The chi square function of EZAnalyze allows you to conduct a "non-parametric" statistical hypothesis test to see if observed frequencies differ from expected values significantly. To keep the EZAnalyze functions consistent, the chi square function requires that the data be "raw" data for analysis. This is different than how chi square is typically used, where you create a chi square table first. The good news is that EZAnalyze creates the chi square table for you from your raw data, so you can interpret the results exactly the same as you would with any other chi square calculator.

To conduct a chi square analysis, you need to identify two categorical variables. For example, if you wanted to know if the number of male and female nurses is significantly different than the number of male and female police officers, your categorical variables would be "gender" and "occupation," with each containing two categories (male/female, and nurse/police officer). This would create a "two by two" chi square table, and tell you if the observed differences in your data are statistically significant. If you asked the same question, but also included firefighters as another occupation, you would then have a two by three chi square table. In EZAnalyze, there is technically no limit to the number of rows and columns your chi square table can contain, but larger tables are difficult to interpret and often not useful unless you have a very large sample.

To conduct a chi square analysis, select "Advanced" from the EZAnalyze menu in Excel, and then choose "Chi square".

When the "Chi square " dialogue box appears, you will be presented with two lists of the variables in your data sheet. Select a variable from the list under "Rows variable" and a different variable from "Columns variable" to create your chi square table. It is usually better to have more rows than columns, but it is really a matter of personal preference!

When you click OK, a results report will be printed on a separate sheet for your review.

Reading your Results Report - Chi Square

Interpreting the results of the Chi Square function

The results report for your chi square analysis contains several pieces of information to help you interpret your results. These are not discussed at length here, but you are encouraged to consult a statistics textbook or the Internet to learn more!

The EZAnalyze chi square analysis produces a chi square table, chi squared value, and a significance level (P) telling you if the observed frequencies were significantly different than the expected frequencies. If your chi

squared value is statistically significant, it is important to look at the actual cells of the chi square table to accurately interpret the results. The number of cells in your table depends on how many levels the variables you chose for analysis had. For example, if you were looking at gender with two levels (male, female) and occupation with three levels (nurse, firefighter, police officer), your chi square table would have 6 cells.

Each cell in the chi square table tells you what your observed value was (the frequency of occurrence from your data) and what the expected value is (the value you would expect to get if there were no differences between the groups). Cells where the observed value are most different from the expected value are where the groups differed, and were the largest contributors. You can get a "hint" of where to look first by looking at the numbers above the "Chi square total" number in the lower right side of your chi square table. These numbers are the "partial chi square" results from the rows, allowing you to see which row contributed the most to the total chi square value.

At the bottom of your chi square table, there is a statement "in English" telling you what your chi square value is, and whether or not it is statistically significant.

Regression

Using the Regression Function

The Regression function in EZAnalyze allows you to conduct a statistical hypothesis test of the ability of a single predictor (or set of predictors) to predict scores on a dependent variable (criterion). Since this is an Advanced feature of EZAnalyze, it will not be discussed at length here. Feel free to consult a statistics textbook or the Internet for information on what Regression is and how it works. If you are already familiar with Regression and want EZAnalyze to crunch the numbers for you, read on!

NOTE: This is the only procedure in EZAnalyze which relies on the 'Analysis Toolpak - VBA', provided by Microsoft. Not all Excel installations contain this program. If your installation does not have the Toolpak, you will not see the Regression option on the EZAnalyze Advanced menu.

To perform a Regression on your data, select "Advanced" from the EZAnalyze menu in Excel, and then choose "Regression".

When the "Regression" dialog box appears, select a criterion (dependent) variable from the list on the left, and your predictor(s) from the list on the right. When you are satisfied with your choices, click OK.

When you click OK, a results report will be printed on a separate sheet for your review.

Reading your Results Report - Regression

Interpreting the results of the Regression Function

The results report for your Regression contains several pieces of information to help you interpret your results. These are not discussed at length here, but you are encouraged to consult a statistics textbook or the Internet to learn more!

The EZAnalyze Regression report contains output generated largely by the Analysis Toolpak, with the only plot being a scatterplot of your regression model's predicted vs. expected values for your criterion variable. These statistics include the Muliple R, R Square, Adjusted R Square, and the Standard Error. You will also receive an ANOVA table, and other standard regression output such as regression coefficients for each of your predictors, significance levels, and the predicted values and residuals for each case in your analysis.

Further discussion of what these are is beyond the scope of this help file!

EZAnalyze also automatically creates a graph of the predicted and expected values for your criterion variable. To learn more about modifying this type of graph, please see *Modifying your Graph*.

The EZAnalyze Manual

OTHER TOOLS IN EZANALYZE

New to Version 3.0, EZAnalyze contains a couple of data management tools that help you prepare data for use in EZAnalyze. These new tools are the 'Amalgamate Data' and "Transpose Data' functions, available from the 'Other Tools' menu in EZAnalyze

Transpose Data

Using the Transpose Data Function

The purpose of the Transpose Data function in EZAnalyze is to take data that violate the 'One Case Per Row' principle of popular data analysis programs and reconfigure the data so that it does not violate the principle, and can be used in EZAnalyze. Many student information systems used in schools construct data in such a way that this principle is violated because it is arguably a better way to put information into storage for easy retrieval (but not statistical analyses). It is important to note that all conversions happen on a copy of your original data sheet, which will remain unharmed during this process.

Since this is a little convoluted, an example may help.

Student ID	Last Name	Date	Grade	Element Description	Score
1234567	Beenthere	8/21/2007	8	Reading Score	2
1234567	Beenthere	8/21/2007	8	Writing Score	3
1234567	Beenthere	8/21/2007	8	Math Score	2
1234568	Seenit	8/21/2007	8	Reading Score	4
1234568	Seenit	8/21/2007	8	Writing Score	3
1234568	Seenit	8/21/2007	8	Math Score	3
1234569	Donethat	8/21/2007	8	Reading Score	1
1234569	Donethat	8/21/2007	8	Writing Score	2
1234569	Donethat	8/21/2007	8	Math Score	3

The data above violates the 'one case per row' principle because each student is actually on three rows; each row with a unique reading, writing or match score.

To use the Transpose Function to reconfigure this data for use with EZAnalyze, select 'Other Tools' from the EZAnalyze main menu, then select 'Transpose Data.

In the 'Transposer' Dialog box, select the variable that contains your Unique Identifier (Student ID in the example above). Next, select the variable that contains a description of the data you are interested in obtaining under the 'Element Description' menu (Element Description is the variable, too, in the example). Finally, select the 'Element Value' from the list provided; this is the actual data you are interested in analyzing (Score in the example above).

OPTION:

- **Delete all other data except these three items** - select this if you want nothing else but the three selected columns (Unique ID, Element Description, Element value) on your data sheet. In the example above, last name, date, and grade would be deleted. You can see in the reconfigured example below what it would look like if this option was not selected.

When you click OK, a new sheet will appear that contains the reconfigured data you selected. If you ran this procedure on the data above, here is what the reconfigured data would look like:

Student ID	Last Name	Date	Grade	Reading Score	Writing Score	Math Score
1234567	Beenthere	8/21/2007	8	2	3	2
1234568	Seenit	8/21/2007	8	4	3	3
1234569	Donethat	8/21/2007	8	1	2	3

Now, these data are ready for EZAnalyze!

The EZAnalyze Manual

Amalgamate Data

Using the Amalgamate Data Function

The Amalgamate Data function in EZAnalyze allows you to combine two worksheets in the same Excel file into one, based on a 'Unique Identifier' that exists in both sheets. For example, if you want to merge data that exist in two different databases (such as NCES and your state Department of Education), you can merge them together easily using this function. This function adds new variables (columns) to your sheet for all cases (rows) that have the same Unique Identifier. You can choose whether or not you want to keep rows of data that do not match.

It is important to note that all additions and deletions happen on copies of your original data, which will not be harmed in any way!

To use the Amalgamate Function, select 'Other Tools' from the EZAnalyze menu in Excel, then choose 'Amalgamate Data'.

In the 'Amalgamator' dialog box, select the worksheets you would like to combine from the drop down menus labeled 'Worksheet 1' and 'Worksheet 2'. Next, select the variable that contains the Unique Identifier in each sheet.

OPTION:

- **Keep matches only** - select this option if you only want to keep cases (rows) that have the unique identifer you selected on both sheets.

When you click OK, your selected data sheets will be combined into one. If any rows of data were added or removed, you will see a 'Case Processing Summary' dialog box informing you of how many additions and deletions there were. Otherwise, you will be brought to the merged data sheet, which is named 'Merged_' plus all or part of the names of the sheets you selected.

Options

Using the EZAnalyze Options

You currently have two options available from this menu in EZAnalyze.

Menu Options. You can choose between either a ‘Cascade Menu’ or the ‘Traditional Menu’. The Cascade menu is the default option for EZAnalyze Version 3.0, while the Traditional menu is available for people who were comfortable with the menu style of previous versions of EZAnalyze.

Uninstall EZAnalyze. You can uninstall EZAnalyze by clicking on the button in the Options dialog box. This will remove EZAnalyze from Excel, and tell you where a file is that can be safely deleted after uninstalling. This file can also be left where it is without harming your computer or using any resources other than a very small amount of disk space.

The EZAnalyze Glossary

ANOVA	Short for "Analysis of Variance," this is used to tell you if at least two subgroups on a categorical variable performed significantly different on a dependent variable. This is one of EZAnalyze's Advanced functions
Area Chart	A graph that has a categorical variable on the X axis and a summary number (mean, total number of people) on the Y axis
Axis - X axis	The bottom of a bar or area chart; shows you the levels of the categorical variable you selected for the graph
Axis - Y axis	The vertical side (left side) of a bar or area chart; displays summary information (mean, total number of people)
Bar Chart	A graph that has a categorical variable on the X axis and a summary number (mean, total number of people) on the Y axis
Case	A row in EZAnalyze. Each case should contribute uniquely to the EZAnalyze data set. For example, if you are analyzing student data, each case would represent one student. If you are analyzing school test scores statewide, each case would be one school
Categorical variable	A variable that divides your data into groups. For example, if you are looking at data for a High school, a good categorical variable would be "grade level." Other examples are gender, group (experimental vs. control) and race
Correlation	A number between negative 1 and positive 1 that indicates the degree of relationship between two variables. The direction of the relationship is indicated by the number being negative or positive, while the strength of the relationship is indicated by the number itself. For example, -.88 would be a strong, negative correlation
Dependent variable	A numeric variable that contains information you are interested in. Test scores, number of days absent, and number of behavioral referrals are good examples of dependent variables
Descriptive statistics	Descriptive statistics are a broad class of statistics used to simply describe your results.
Difference score	A difference score is a new variable that is the difference between two other variables. Allows you to show changes over time with one variable
Disaggregate	Disaggregate means to sort something into categories. In EZAnalyze, disaggregate means to sort your dependent variable by your categorical variable
Distribution of scores	The distribution of scores is the range of scores people obtained on a variable, and how many people scored in each category of the variable. For example, the distribution of scores for a variable called gender could be "65 males and 45 females"
EZAnalyze	A fun software program that makes data analysis a snap!
Frequency	The frequency is a simple count of how many times something occurred. For example, if 55 people said "yes" to a question, the frequency would be 55
Maximum value	The highest number observed in your data for the given variable
Mean	The average score of all of the scores in a given range of values
Median	The middle occurring number if you laid all of your scores out on a line. For example, the median of 0, 0, 1, 2, 3 is 1, and the median of 5, 5, 5, 5, 100, 111, 112, 113, 114 is 100
Minimum value	The lowest number observed in your data for the given variable
Missing value	A number that is not present in your data, and is not included in analyses
Mode	The most frequently occurring number. For example, 3 is the mode of 0, 1, 2, 3, 3 because there are more 3's than any other number
NTV	A "Numeric Test Value." Used to specify a known mean in the one sample t-test
Numeric variable	A variable that contains "meaningful numbers". For example, if you use the numbers 1 and 2 to represent males and females in your data, "gender" would not be a numeric variable (even though it contains numbers). If you can get a mean score for the variable that makes sense, then it is probably numeric. For example, saying that "the average race in our data was Caucasian" does not make sense. Saying "the average test score was 98" does make sense
Percent	Mathematically, the number in a category divided by the total number in all categories. Simply put, it is the proportion of scores on a variable relative to the total number of scores
Percentages	See *Percent*
Pie Chart	A chart that displays the percent of people in each category. Each slice of the pie represents a category, and the percent of the total that each slice represents is its size
Results report	A chart or table that displays the results of an EZAnalyze function
Standard deviation	How far, on average, each score in your data deviates from the mean. The larger the standard deviation, the farther each person's score on average was different than the mean
String variable	A variable that contains letters. Can not be used as a dependent variable
T-Test, independent	A statistical hypothesis test used to indicate the degree of difference between two group means. An EZAnalyze Advanced function
T-Test, one sample	A statistical hypothesis test used to indicate the degree of difference between an observed sample mean and the value of a known population mean. An EZAnalyze Advanced function
T-Test, paired	A statistical hypothesis test used to indicate the degree of difference between two non-independent (paired) means. An EZAnalyze advanced function
Total score	The total of all selected data pieces added up
Valid N	The number of people who had usable data for the selected analysis
Variable	A column in EZAnalyze. More specifically, a variable is something that can assume different values. The variable "gender" can assume two values, while the variable "gross income" could assume a potentially infinite number of values
X axis	The bottom of a bar or area chart; shows you the levels of the categorical variable you selected for the graph
Y axis	The vertical side (left side) of a bar or area chart; displays summary information (mean, total number of people)