Increasing Graphing Literacy and Graphing Ability in Undergraduate Psychology Majors

Through Active Learning Based Exercises

Meridith Pease Selden

Yuba College*

Supported by a 2010 Instructional Research Award

* A substantial portion of the grant was completed at Wilkes University

Author contact information:
Meridith Pease Selden
Department of Psychology
Yuba College
2088 N Beale Rd,
Marysville, CA 95901
mselden@yccd.edu
(530) 741-6854

Copyright 2014 by Dr. Meridith Pease Selden. All rights reserved. You may reproduce multiple copies of this material for your own personal use, including use in your classes and/or sharing with individual colleagues as long as the author’s name and institution and the Office of Teaching Resources in Psychology heading or other identifying information appear on the copied document. No other permission is implied or granted to print, copy, reproduce, or distribute additional copies of this material. Anyone who wishes to produce copies for purposes other than those specified above must obtain the permission of the author(s).
TABLE OF CONTENTS

Overview and Explanation

Types of Data ...........................................................................................................4
Research Methods & Variables .................................................................................6
Example Bar Graph (Figure 1) ..................................................................................10
Example Histogram (Figure 2) .................................................................................11
Example Frequency Polygon (Figure 3) .................................................................12
Example Scatterplot (Figure 4) ................................................................................13
Example Line Graph (Figure 5) ..............................................................................14

Activities

Types of Data Questions ..........................................................................................15
Fix the Graph ............................................................................................................17
Graph Questions I .....................................................................................................18
Graph Questions II ....................................................................................................19
M&Ms Activity ..........................................................................................................20
Wadded Up “Paperball” Toss Activity ....................................................................21
Survey Activity ..........................................................................................................22
  Survey .....................................................................................................................23
  Survey (Types of Data) Questions ..........................................................................24

“How to” guides (Microsoft Excel*)

Bar Graph ...............................................................................................................26
Histogram ...............................................................................................................30
Frequency Polygon ...............................................................................................38
Scatterplot ...............................................................................................................50
This resource contains a series of activities designed to increase graphing literacy and graphing ability in undergraduate psychology majors. The activities involve an activity based approach to graphing, emphasizing the importance of being able to read/understand graphs and providing detailed instructions (including screen shots) of how to create a variety of types of graphs in Excel 2007 (v12.0).

This resource should be relevant to all psychology students and educators. However, it is primarily aimed at faculty who teach statistics and/or research methods. More specifically, all of the activities are designed to be completed in class thereby encouraging their inclusion in the graphing unit of a statistics and/or research methods class.

Outside of a statistics and/or research methods class, any instructor that teaches a class involving student consumption of journal articles (or other primary sources) might find the resource relevant as it contains information not only on how to create graphs, but also how to understand them. Additionally, each activity was designed as “stand alone” activity so they can be incorporated as individual activities in any class that deals with graphs or graphing.
Types of Data

Qualitative and Quantitative Data

**Qualitative Data.** Classified based on characteristics and not measured in amounts.

**Quantitative Data.** Measured in and indicates amounts (e.g., number present).

The graphic below illustrates how data are classified after being identified as qualitative or quantitative.

Four Traditional Measurement Scales

**Nominal scale.** Scores are used for identification and classification with no quantitative properties. Each score represents belonging to a group, not an amount. For example, if a class were split into males and females, this would be a nominal scale.

The nominal scale represents qualitative data and is the most basic measurement scale. It is a discrete measurement meaning that membership in one category precludes membership in another group. For ease of understanding, nominal scales often assign numbers to the categories (e.g., Males = 1, Females = 2). However, the number has no inherent value associated with it and is only used for identification so a female (assigned a 2) is not any better or worse than a male (assigned a 1).

With nominal data, a **bar graph** is the best way to show the data.
**Ordinal.** Scores indicate rank order. Categories are more than merely different and can be rank ordered from highest to lowest or from best to worst. For example, if a class were split according to grade (i.e., A, B, C, D, F), this would be an ordinal scale. With ordinal scales, like with nominal scales, membership in one category precludes membership in another group.

There is some disagreement about whether ordinal data represent qualitative or quantitative data. Although ordinal data can be ranked, only the respective rankings can be discussed; nothing can be said about data between the ranks. That is to say, it is known that an A is higher than a B but it cannot be known by how much. It is possible the A (e.g., 90%) may only differ from the B (e.g., 89%) by one point. However, it is equally possible that the A (e.g., 99%) differs from the B (e.g., 80%) by 19 points.

With ordinal data, a **bar graph** is the best way to show the data.

**Interval/Ratio.** Interval scales have scores that represent an actual quantity with an equal distance between the scores (e.g., 100 is 10 points higher than 90 and 90 is 10 points higher than 80). Interval scores do not contain a true zero and thus do not measure the complete absence of a quantity. For example, if a class were organized based on IQ score, this would be an interval scale. There is not a true zero here because one cannot score a zero (indicating a complete lack of intelligence) on the IQ test and we cannot say that someone who scores a 100 is twice as smart as someone who scored a 50.

Ratio scales, on the other hand, also have scores that represent an actual quantity with an equal distance between the scores. However, ratio scales contain a true zero. For example, if a class were organized based on the time it took to complete the first exam, this would be a ratio scale. There is a true zero here because it is possible to take no time to complete the exam (e.g.,
the student did not attempt the exam at all), and a student who spends 60 minutes takes twice as long as a student who spends 30 minutes to complete the exam.

The appropriate way to present data and the statistical tests are identical for interval and ratio scales. Additionally, both interval and ratio scales contain continuous data. Thus, the scores can have decimal places (e.g., 90.5), indicating that a score falls between two whole numbers. As such, with Interval/Ratio data, either a histogram or a frequency polygon, both of which show the continuous nature of the data, is the best way to show the data. A histogram, which is appropriate when there are fewer scores, contains bars that abut one another. A frequency polygon, which is appropriate when there are a larger number of scores, contains an unbroken line.

It is worth noting one special case of interval and ratio scales in which the data are discrete. Namely, when one presents counts (e.g., the number of babies born a year) the data are discrete but also ratio. In this case, even though the data cannot technically be a decimal (e.g., 2.3 babies), a histogram or a frequency polygon is still the appropriate way to show the data.

**Research Methods and Variables**

Psychological research commonly uses three methods: descriptive, correlational, and experimental.

**Descriptive**

Of the three commonly used types of research, descriptive research is the most basic. Descriptive research does exactly what it sounds like; it describes a data set. All of the information presented in descriptive research can be directly verified by the data. In other words, no inferences are drawn about other data or other situations.
All of the following (made up) examples represent descriptive research: The average statistics grade on the first exam was 71.43, two thirds of the students enrolled in general psychology are women, and only one first-year student is enrolled in research methods. As you can see, all of these examples describe something about a group of people and can be verified by examining the data. None of the examples make claims about other groups or do anything beyond conveying basic information about the data set.

With descriptive research, three types of graphs are typically used: bar graphs, histograms, and frequency polygons.

**Correlational**

Correlational research is sometimes classified as descriptive and sometimes classified as its own separate category. This ambiguity arises because a correlation can be used to merely describe the strength of a relationship or it can be used to test for the presence of a relationship. In either situation, the purpose of correlational research is to examine the relationship between two variables. The difference between descriptive research and correlational research is that correlational research emphasizes relationships whereas descriptive research does not.

All of the following (made up) examples represent correlational research: There is a positive relationship between the number of hours studied and the grade earned on the first statistics exam, increased amounts of estrogen are related to higher levels of empathy, and the amount of beer a person consumes is negatively related to the person’s ability to drive well.

With correlational research, only one type of graph is typically used: a scatterplot.

**Experimental**

Experimental research is the most complicated of the three types of research. Consequently, it provides the richest data. Experimental research separates itself from other
types of research in that the experimenter has control over one (or more) variable(s). For example, if researchers want to examine whether consuming caffeine makes someone more alert, they will need to conduct an experiment. One potential (although simplistic) way of examining this effect would be to give participants different amounts of caffeine (e.g., 50 mg, 100 mg, and 150 mg) in a between-subjects design and test for alertness. In this case, the experimenters control the amount of caffeine a person consumes and it is this control that makes the study an experiment. Additionally, because the experimenters are giving different amounts of caffeine to different participants, they are manipulating the amount of caffeine received. Manipulation is a key component of experimental research. Although experiments are powerful and frequently used, they may be less realistic because the experimenter controls and manipulates the situation.

Experimental research has two types of variables unique to experiments. One is the variable that is under the control of and deliberately manipulated by the experimenter. This variable is called the independent variable (IV). The independent variable will always have at least two discrete (nonoverlapping) levels. In the example described above, the experimenter created three levels: 50 mg, 100 mg, and 150 mg of caffeine. Often, experiments will consist of more than one independent variable. Multiple independent variables allow for the study of complex human behavior. For example, with the caffeine study discussed above, it is entirely possible that the time of day the caffeine is consumed may have an effect on a person’s ability to focus. Therefore, the experiment can create two levels corresponding to the time of day the caffeine is consumed: morning (10:00 am) or afternoon (3:00 pm). The experimenter would assign participants to either the morning group or the afternoon group, thereby manipulating the time of day the caffeine is consumed.
The second type of variable that all experimental research has is called the dependent variable (DV). The dependent variable is the variable being measured and is not controlled by the researcher. The DV is the data (D) the researcher collects. In the above example, the researcher measures each participant’s level of alertness so the level of alertness is the DV. Another way of thinking of a DV is that it depends on the IV. That is to say, the person’s level of alertness depends on the amount of caffeine consumed (assuming caffeine has an effect on alertness). As opposed to the IV, the DV in this example is continuous (the opposite of discrete data and includes all data that have the potential to include fractions or decimals).

A final quality of experimental research is that it is often inferential. That means it is used to make inferences about other similar situations and groups of people. Although it is interesting to know that caffeine has an effect on the level of alertness for a particular sample, the goal is to be able to apply that knowledge in other situations. In other words, the goal of the study is to be able to draw the inference that in general caffeine increases alertness, not just in the specific situation studied.

With experimental research, two types of graphs are typically used: bar graphs and line graphs.
Figure 1. Example of a bar graph.
There should be a descriptive title for the graph.

Scores on the First Statistics Exam

There should be a descriptive title for the Y-axis.

There should be a descriptive title for the X-axis.

The bars should be touching.

The height of each bar represents the frequency with which each category occurs in the data set.

Each category on the X-axis should be labeled.

The X-axis labels should be centered under the axis tick marks (indicating the upper and lower real limits of the continuous data).

Number of Students

Statistics Exam Scores

Figure 2. Example of a Histogram.

There should be a descriptive title for the graph of the first statistics exam scores. The bars should be touching, and each category on the X-axis should be labeled. The X-axis labels should be centered under the axis tick marks (indicating the upper and lower real limits of the continuous data). The height of each bar represents the frequency with which each category occurs in the data set.
Figure 3. Example of a frequency polygon.

Scores on the First Statistics Exam

Number of Students

Statistics Exam Score

The line needs to connect to the X-axis before the first interval and after the last interval.

The X-axis labels should be centered under the axis tic marks (indicating the upper and lower real limits of the continuous data).

Each category on the X-axis should be labeled with the value in the middle of each interval – 73 is the middle of the 70 to 76 interval. If you are not graphing intervals then the X-axis should be labeled with whole numbers.

The height of each bar represents the frequency with which each category occurs in the data set.

There should be a descriptive title for the graph.

There should be a descriptive title for the graph.

There should be a descriptive title for the Y-axis.

The two // indicate that the X-axis does not start at 0. The lines can be added to a graph by manually adding a text box. However, for the purposes of this demonstration, it is easiest to draw the lines in after the graph has been printed.

There should be a descriptive title for the X-axis.

There should be a descriptive title for the graph.

The X-axis does not start at 0.
**Figure 4.** Example of a scatterplot.
There should be descriptive title for the graph

There should be a descriptive title for the Y-axis

Scores on a Concentration and Focus Skills Test as a Function of the Amount of Time of Caffeine Consumption

Scores on the Concentration and Focus Skills Test

Amount of Caffeine Consumed

Morning (10:00 am)

Afternoon (3:00 pm)

Each level on the X-axis should be labeled

The lines should be easy to see and to differentiate

There should be a legend that clearly labels each line on the graph

Unless indicated by a break in the axis (i.e., //), each axis should start at zero

Figure 5. Example of a line graph.
Types of Data Questions

1. Social psychologists often use nominal variables in their research. What are some nominal variables you would expect a social psychologist to regularly use?

2. Cognitive psychologists often use interval/ratio variables in their research. What are some interval/ratio variables you would expect a cognitive psychologist to regularly use?

3. Provide an example of a research scenario where a psychologist would collect ordinal data.

   a. Do you think rank data are better classified as quantitative or qualitative? Why?

4. In social science research, which do you think you will encounter more as a dependent variable: nominal, ordinal, or interval/ratio data? Why?
5. One controversy among social science researchers regards multiple-choice measures such as Likert-type questions (e.g., 1 = Strongly Disagree, 2 = Disagree, 3 = Neither Agree nor Disagree, 4 = Agree, 5 = Strongly Agree). Some researchers consider these types of questions to be ordinal and some consider them to be interval. Which do you consider them to be, ordinal or interval? Why?

a. What are the possible ramifications of classifying the measure as ordinal?

b. What are the possible ramifications of classifying the measure as interval?
Fix the Graph

Below are two flawed graphs. Determine what is wrong with each graph and fix it. If you need to draw a new graph, you can.

**Age of Students Enrolled in Statistics**

- The x-axis labels do not represent the correct age groups.
- The y-axis should represent the number of students instead of the scale from 0 to 7.
- The bars should be clearly labeled with the age groups.

**Student Location**

- The y-axis label is not clear, it might be representing the number of students.
- The bars for Off-Campus and On-Campus should be clearly separated and properly labeled.
Graph Questions 1

Use the graph below to answer the following questions:

![Graph showing percent of participants who helped a person in need as a function of sex and status]

1. What kind of graph is it?

2. What is the Dependent Variable (DV)?

3. How many Independent Variables (IVs) are there?
   a. How can you tell?
   b. What are they?
4. When are people the most likely to help, when they are alone or in a group?

Graph Questions 2

Use the graph below to answer the following questions:

![Bar graph showing quiz scores]

1. What kind of graph is it?

2. What kind of data (nominal, ordinal, interval, or ratio) were used to make the graph? How do you know?

3. Are there any other graph types that would be appropriate to use? If so, what kind?

4. How many students scored a 6?

5. Which score was the most frequently occurring score in the class?

6. How many students are enrolled in the class?
**M&Ms Activity**

For this activity, you will need a pack of M&Ms and access to a computer with Microsoft Excel. Using the M&Ms, you will need to do the following:

1. Divide the candies by color.

2. Count the number of each color.

3. Create a graph in Microsoft Excel** that shows the breakdown by color.
   a. Make sure to select the most appropriate type of graph.
   b. Don’t forget to include labels and a descriptive title!

*NOTE: Although this activity was not taken from a specific source, there are many versions of it in use by a variety of people. However, none of the existing versions were consulting when making this activity.

**See the “How-to” guides at the end of this workbook for step-by-step instructions on how to make each of the five main types of graphs (i.e., bar, histogram, frequency polygon, scatterplot, and line graph)
**Wadded Up “Paperball” Toss Activity**

For this activity, you will need a sheet of paper, a yardstick or tape measure, and access to a computer with Microsoft Excel. Using the sheet of paper and the ruler, you will need to do the following:

1. Wad up a sheet of paper to make a “paperball.”
2. Mark a spot on the floor from which everyone will throw their paperball.
3. Each person in the class should throw a paperball and then measure the number of inches it flew.
   a. To ensure consistency in the measurement, each person should measure from the tips of their toes (at the time of the toss) to the final resting place of the “paperball”
      i. Try to measure the most direct route possible.
   b. All measurements should be in inches
4. Create a graph in Microsoft Excel* that shows the paperball distances
   a. Make sure to select the most appropriate type of graph.
   b. Don’t forget to include labels and a descriptive title!

*See the “How-to” guides at the end of this workbook for step-by-step instructions on how to make each of the five main types of graphs (i.e., bar, histogram, frequency polygon, scatterplot, and line graph)
Survey Activity

For this activity, you will need the survey that is included with this workbook on the next page and access to a computer with Microsoft Excel. You will need to do the following:

1. Have each person in the class complete a survey.
2. Collect and enter the data into Microsoft Excel*.
3. Select a question that interests you and that uses **nominal** data.
   a. Create the appropriate type of graph in Microsoft Excel.
   b. Don’t forget to include labels and a descriptive title!
4. Select a question that interests you and that use **ordinal** data.
   a. Create the appropriate type of graph in Microsoft Excel.
   b. Don’t forget to include labels and a descriptive title!
5. Select a question that interests you and that uses **interval/ratio** data.
   a. Create the appropriate type of graph in Microsoft Excel.
   b. Don’t forget to include labels and a descriptive title!

*See the “How-to” guides at the end of this workbook for step-by-step instructions on how to make each of the five main types of graphs (i.e., bar, histogram, frequency polygon, scatterplot, and line graph)
1. How old (in years) are you? ____________________

2. Are you working? Yes
   Full-time
   Yes
   Part-time
   No
   If yes, how many hours a week do you work?
   (Provide a number from 0 to 40) ____________________

3. What is your sex? Male
   Female

4. Student Status? Full-time
   Part-time

5. What is your current residence as a student? On-campus
   Off-campus

6. What year are you in school?
   Freshman  Sophomore  Junior  Senior

7. Are you a psychology major? Yes
   No

8. What percentage do you hope to earn in this class?
   (Provide a percent from 0% to 100%) ____________________

9. How many hours a week do you plan to study for this class?
   (Provide a number from 0 to 40) ____________________

10. Approximately, how many miles away from campus did you grow up?
    0 to 20 miles  61 to 80 miles
    21 to 40 miles  81 to 100 miles
    41 to 60 miles  More than 100 miles

11. What do you think is the best thing about your university?

12. What do you think is the worst thing about your university?
Survey (Types of Data) Questions

1. Which of the questions on the survey resulted in nominal data?

   a. Sketch an example of the type of graph that should be used to show the data.

2. Which of the questions on the survey resulted in ordinal data?

   a. Sketch an example of the type of graph that should be used to show the data.

3. Which of the questions on the survey resulted in ratio data?

   a. Sketch an example of the two types of graphs that should be used to show the data.
4. In addition to the survey’s nominal questions, the open-ended questions at the end are also qualitative questions. How do you think the data collected from these qualitative questions will differ from the data collected from the other questions?

a. Which type of data do you think you would prefer to work with, qualitative or quantitative? Why?
How to Create a Bar Graph

- There should be a descriptive title for the graph.
- There should be a descriptive title for the Y-axis.
- Unless indicated by a break in the axis (i.e., //), each axis should start at zero.

- The bars should not be touching.
- The height of each bar represents the frequency with which each category occurs in the data set.
- Each category on the X-axis should be labeled.

**Undergraduate Student Enrollment in the College of Social and Behavioral Sciences**

- Number of Students
- Undergraduate Major
- Psychology: 35
- Anthropology: 5
- Sociology: 15
- Criminology: 20
- Economics: 10
- International Studies: 5
- Political Science: 10
These counts represent the number of times each major occurs in the data set.

Highlight the counts and the labels.

Click the "Column" Icon.

Click the on the first icon in the "2-D Column" Row.

Click the "Insert" tab.
INCREASING GRAPHING LITERACY AND GRAPHING ABILITY

Typically, the Independent Variable (IV) goes on the X-axis and the Dependent Variable (DV) goes on the Y-axis.

With the graph highlighted, select the bottom arrow under “Chart Layouts.”

Selecting the bottom arrow will provide a drop down menu. Select the graph preview image that contains spaces for a title and both axes. (In this case, that image is the ninth image.)

Add a descriptive title and label both axes.

HINT: The Y-axis contains the counts and the X-axis contains the variable in which you are interested.

The graph needs a title and the axes need to be labeled.
To remove the (unnecessary) grid lines, select one line and right click.

Select “Delete.” This will delete all of the gridlines.

Select “Delete.”

To remove the (unnecessary) key, select “Series1” and right click.
How to Create a Histogram

There should be a descriptive title for the graph.

Scores on the First Statistics Exam

The bars should be touching.

There should be a descriptive title for the Y-axis.

The height of each bar represents the frequency with which each category occurs in the data set.

Unless indicated by a break in the axis (i.e., //), each axis should start at zero.

The X-axis labels should be centered under the axis tick marks (indicating the upper and lower real limits of the continuous data).

There should be a descriptive title for the X-axis.

Each category on the X-axis should be labeled.
You will need to **gather the frequencies of each score**. To do this you will need to write a “countif” function for each score. The “countif” function tells Excel to count all of the specified values in a specified range. Since there are 11 scores (0, 10, 20, 30, 40, 50, 60, 70, 80, 90, & 100) you will write 11 “countif” functions.

This function can either be written by clicking on any empty cell or by clicking on an empty cell and writing it into the function (fx) bar.

For this example, the range is A1:B21 (indicating that excel is to look at cells 1 to 21 in column A and 1 to 21 in column B).

Before you gather the frequencies, you will want to set up the labels – The labels should contain all of the scores in the data set, starting with the largest and ending with the smallest.

Every function MUST start with an “=” sign.

Note the “ ” around the 0 – The “ ” tells excel which value to count. If you forget to add the “ ”, you will get an error message.

After you type in the function (Don’t forget the = and “!”) hit “Return/Enter”.

There should be a number in the cell representing the frequency count of the specified value.
The frequency distribution for the dataset should end up looking like this:

Notice the blank cell before the “0” and after the “100” – This is necessary for formatting reasons.
You will need to add the blank spaces above the lowest score and below the highest score.

These are example “countif” functions – When you write the functions they will not show up like this (you should only get the numbers).

Notice the blank cell before the “0” and after the “100” – This is necessary for formatting reasons.
You will need to add the blank spaces above the lowest score and below the highest score.
Highlight the frequency values that you just calculated.

Click the "Insert" tab.

Click the "Column" icon.

Clicking the "Column" icon will cause the following drop down menu to appear – Select the icon under "2-D Column" that shows 2 sets of 2 grouped bars (the first icon in this picture).

Highlight the frequency values that you just calculated.
This is what your graph should initially look like
Note: It is currently NOT a histogram (it’s a bar graph)

To turn the bar graph into a histogram and to add a title and axes label, you need to change the chart layout
Click the bottom most arrow in the chart layout area

Select the icon that shows solid bars touching. In this case, that icon is the eight icon

Your graph should now look like this
To add appropriate X axis labels AND to put tick marks on the X axis, right click while the axis number (1, 2, 3, etc.) is highlighted.

Double click on “Axis Title” to add a descriptive Y axis title.

To center the tick marks, click on “Format Axis”.

To add appropriate axis labels click on “Select Data”.

Double click on “Axis Title” to add a descriptive X axis title.

Double click on “Chart Title” to add a descriptive title.
Select the "On tick marks" button.
This centers the tick marks under the axis labels.

To change the labels, click "Edit".

This is a box (which automatically pops up when you click edit) allows you to tell Excel where the labels are located. It is blank because no labels have been identified.

This is telling you where the data are located. This is correct so you will not need to make any changes here.
This is blank because no labels have been identified.
To identify the labels, highlight the labels in the dataset.

For this example, the labels are in column D.

Make sure to include the blank cells you inserted in the frequency distribution, they are needed for formatting reasons.

After you have highlighted the labels in the dataset, click “OK”.

Scores on the First Statistics Exam
How to Create a Frequency Polygon

Scores on the First Statistics Exam

The height of each bar represents the frequency with which each category occurs in the data set.

The line needs to connect to the X-axis before the first interval and after the last interval.

The X-axis labels should be centered under the axis tic marks (indicating the upper and lower real limits of the continuous data).

Each category on the X-axis should be labeled with the value in the middle of each interval – 73 is the middle of the 70 to 76 interval. If you are not graphing intervals then the X-axis should be labeled with whole numbers.

There should be a descriptive title for the graph.

There should be a descriptive title for the Y-axis.

The two // indicate that the X-axis does not start at 0. The lines can be added to a graph by manually adding a text box. However, for the purposes of this demonstration, it is easiest to draw the lines in after the graph has been printed.

There should be a descriptive title for the X-axis.

The X-axis labels should be centered under the axis tic marks (indicating the upper and lower real limits of the continuous data).
Before creating the graph, you will need to **create intervals**. Intervals group the data into manageable categories. A general rule of thumb is to have approximately 10 intervals per data set. Therefore, the easiest way to determine the appropriate interval size for a data set is to take the range (i.e., highest score – lowest score) of the data set and divide it by 10. (NOTE: It is also possible to create a polygon using ungrouped data. If you are using ungrouped data you can skip following grouping steps.)

After determining the interval size, you will need to write a “countif” function for each interval. The “countif” function tells Excel to count all of the specified values in a specified range. You will need to specify each value in the interval that is to be counted. Since there are 10 intervals you will write 10 “countif” functions.

### Data Set

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>29</td>
<td>74</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>42</td>
<td>74</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>43</td>
<td>75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>52</td>
<td>75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>54</td>
<td>76</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>55</td>
<td>76</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>56</td>
<td>77</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>58</td>
<td>77</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>59</td>
<td>79</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>60</td>
<td>81</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>62</td>
<td>82</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>63</td>
<td>83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>65</td>
<td>84</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>66</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>68</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>71</td>
<td>87</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>71</td>
<td>91</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>71</td>
<td>97</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>74</td>
<td>97</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This is the data set. It is in two columns to make it easier to see. However, you will likely have it in only one column. The number of columns used does not change the process.

The “(A1:B20)” indicates to excel where the data is located. This corresponds to the row and column labels in the upper left corner around the data set.

As described above, the first step is determine the **range** of the data – This text tells Excel to find the largest and smallest number in the data set and then to subtract the smallest number from the largest number.

When you type this into Excel, be sure to include the “=” sign. Without an “=” sign, Excel will not treat the text as a function.

The range is 68
The range is 68. As described above, the general rule of thumb is 10 intervals per data set. So, if you divide the range by ten (68/10=6.8) you will get the interval size.

NOTE: round to the nearest whole number.

Once you have determined the interval size, the next step is to create your intervals. This will involve a few steps. However, the first step is to list the intervals. A few tips (Pagano, 2007):

- The width of the interval is 7 units
- Formatting purposes, start with the smallest interval and work to the largest interval
- The lower limit of the lowest interval must contain the lowest score in the data set
- The lower limit of the lowest interval (i.e., 28) should be evenly divisible by the interval size (i.e., 7)
- Although it may seem like the intervals are only 6 units wide, they are indeed 7 units (i.e., 28, 29, 30, 31, 32, 33, and 34)
- There are blank cells above the smallest interval and below the largest interval for formatting reasons
- It is easiest to just type the intervals into Excel. In this case, you want to be sure to NOT include an “=” sign.
Now that you have labeled the intervals, the next step is to get the frequencies within each interval. In other words, using the Excel “countif” function, you will be counting the number of times each value within the interval occurs in the data set. So, for the lowest interval (28-34), you will be counting the frequency with which each value in the interval (28, 29, 30, 31, 32, 33, and 34) occur in the data set. In this case, the only value that occurs is 29 and it only occurs one time so the frequency is 1.

Although this looks like a long formula, it is not that complicated. There are 7 “countif” statements (one for each unit in the interval) embedded in the “sum” function. The “sum” function tells Excel to add the frequencies for each unit within the interval. (The data for this example are organized by size for demonstration purposes. Your data do not need to be organized in this manner.)

An individual “countif” statement looks like this:
COUNTIF(Data Range, Value) where the data range corresponds to the column and row labels surrounding the data set (in this case it is A1 to B20 which indicates that the data set starts in cell A1 and finishes with B20) and the value corresponds to the specific number within the interval you are counting

Each “countif” statement needs to be separated by a comma

The “sum” command needs to be preceded by an “=” sign AND needs to be followed by “( )”. ALL of the “countif” statements should be contained within the “( )”.

```
```
This is what all of the “countif” statements look like for this example.

If the count if statements are too complicated/confusing, you can always count the number of values within each interval by hand and then enter the counts into excel. However, if you decide to do this, count carefully!!

TIP: The sum of the f column should equal the total number of data points in your data set. If it does not then you’ve made an error.

These are the values Excel provides using the above “countif” statements. Note that $1 + 0 + 2 + 3 + 6 + 2 + 12 + 7 + 3 + 4 = 40$ and there are 40 people in the data set.

If you chose to count by hand, you should have the same frequencies for each interval.
Now that you have the intervals and the frequencies, you need to add labels that will be used on the graph. These are different than the interval labels. For the graph, you will label each interval with the value in the middle/center of the interval. So, for the 28-34, the middle value is 31.

28, 29, 30, 31, 32, 33, 34

HINT: After you determine the middle of the lowest interval, you can add the width of the interval (i.e., 7) to the value to get the next label.

31 + 7 = 38
(38 is the middle of the second interval)

The 0 above the lowest interval and below the highest interval is included for formatting reasons (it indicate that there are values that do not occur in the data set). You will also need to include a label for the graph (again for formatting reasons). This will be explained below.

For formatting reasons, include a label below the smallest interval (i.e., 24) and above the largest interval (i.e., 101).

These are all of the labels for the intervals.
Now you’re finally ready to create the graph 😊

Click “Insert” in the menu bar.

Under the “Line” option, select the first 2-D line option. It shows two lines in the preview but will have only one line in the graph.

Highlight the frequency column. Be sure to include the leading zeros you added on the previous page.

Following steps 1 - 3 will yield this graph. As you can see, it is a good start but it is not formatted or labeled correctly.
INCREASING GRAPHING LITERACY AND GRAPHING ABILITY

1. To add a title, while the graph is selected, click on "Layout" under chart tools and then select "Chart Title" and "Above Chart".

2. Add the title where it says "Chart Title."

3. To label the X-axis, while the graph is selected, click on "Layout" under chart tools and then select "Primary Horizontal Axis Title" and "Title below the Axis"

4. Add the X-axis label where it says "Axis Title."

5. Repeat the same process (steps 3 & 4) for the Y (vertical) Axis, selecting "Rotated Title" instead of "Title below the Axis".
To remove the (unnecessary) legend, right click.

To remove the (unnecessary) grid lines, double click directly ON one of the lines.

Select “Delete” twice to remove both the legend and grid lines.
The final thing that needs to be done is to label the X-axis. You will use the label you created previously.

1. Right click anywhere in the graph.
2. Select “Select Data”.
3. You will be changing the Axis label only. Click “Edit” under the Horizontal (Category) Axis Labels.
Select the blue and red box to the right of the "Axis label range:"

After selecting the blue and red box in step one, highlight the graph labels you created earlier.

NOTE: You will leave the "Axis Labels" dialog box to do so. The labels are on the main Excel spreadsheet.

After highlighting the graph labels, click the blue and red box.

Click "OK"
Click “OK”
How to Create a Scatterplot

There should be a descriptive title for the graph.

There should be a descriptive title for the Y-axis.

The predicted variable (i.e., typically the variable of interest) goes on the Y-axis.

Unless indicated by a break in the axis (i.e., ||), each axis should start at zero.

There should be a descriptive title for the X-axis.

The predictor variable (i.e., typically the "known" variable) goes on the X-axis.

Categories on the X-axis should be labeled.

Each point represents a participant’s score on BOTH tests.

Scatterplot of the Relationship Between Scores on Exam 1 and Exam 2.

Each point represents a participant’s score on BOTH tests.
INCREASING GRAPHING LITERACY AND GRAPHING ABILITY

This is the data for the graph. The order of the data does not matter but the data needs to be organized by PERSON.

Highlight the two columns of data.

With the data still highlighted, click the “Insert” tab.

Under “Scatter,” select the image that shows individual dots (in this case, the first image).

This is a good start but some things need to be added (i.e., a title, etc.)
With the chart highlighted, select the “Layout” tab under “Chart tools”.

After you select “Chart tools” you will have the option to change the chart layout.

Select the bottom arrow.

Selecting the bottom arrow will bring up 11 different chart layouts. Select the icon that has a title, axes labeled, and a key (in this case, the first layout option).
Add a descriptive title
Label both axes

Excel has many options for colors, shapes, and sizes. For an APA style paper, be sure to select the gray/black option (indicated here). For a poster or presentation, you may wish to select another color. For this document a neutral blue was used. A regular (non-color) printer will print the blue dots as light gray.
To remove the grid lines select one of the lines and right click

Select “Delete”

To remove the legend, select it and right click

Select “Delete”
If you want to have a shape that is not the default shape, select a data point in the graph and right click.

Select "Format Data Series".

Go to "Marker Options" and select "Built-in".

Select any shape/size that you like.
How to Create a Line Graph

Scores on a Concentration and Focus Skills Test as a Function of the Amount of Time of Caffeine Consumption

- The lines should be easy to see and to differentiate
- Each level on the X-axis should be labeled
- There should be a legend that clearly labels each line on the graph
- There should be a descriptive title for the graph
- There should be a descriptive title for the Y-axis
- Unless indicated by a break in the axis (i.e., //), each axis should start at zero
- There should be a descriptive title for the X-axis

Scores on the Concentration and Focus Skills Test

Amount of Caffeine Consumed

- Morning (10:00 am)
- Afternoon (3:00 pm)
This is the data for your line graph.

<table>
<thead>
<tr>
<th>Caffeine Levels</th>
<th>Time of Day</th>
<th>Morning (10:00 am)</th>
<th>Afternoon (3:00 pm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 mg</td>
<td>50 mg</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>100 mg</td>
<td>100 mg</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>150 mg</td>
<td>150 mg</td>
<td>12</td>
<td>14</td>
</tr>
</tbody>
</table>

Time of day is one of two Independent Variables (IV). There are two levels to this IV.

Each mean represents a condition. For this experiment, there are 6 conditions:
- 50 mg in the morning
- 50 mg in the afternoon
- 100 mg in the morning
- 100 mg in the afternoon
- 150 mg in the morning
- 150 mg in the afternoon

The numeric values are the mean of the Dependent Variable (DV) for each condition. The next page will demonstrate how to calculate means in Excel.

For this demonstration, the Dependent Variable (DV) is the score on a focus and concentration skills test.

Caffeine level is one of two Independent Variables (IV). There are three levels to this IV (50 mg, 100 mg, and 150 mg).
For this demonstration, we will be calculating one of the six condition means. Specifically, we will be calculating the mean of the focus test performance for participants who ingested 50 mg of caffeine in the morning. You will need to calculate a mean for each condition.

Every function MUST start with an “=” sign. In this case, the function is called “AVERAGE”.

Type =AVERAGE and then highlight or type in the range of data. For this example there are 9 participants all located in the A column so the range is A2:A10.

After you type in the function (Don’t forget the = and “!”) hit “Return/Enter”

There should be a number in the cell (in this case, A11) representing the calculated average.
Click the “Insert” tab

Select the graph icon with the two crossing lines under the “2-D Line” options. In this case, it is the first icon.

Select the cell labels AND the mean values (but not the variable names).

Excel automatically puts in two solid, colored lines. To change the color or format of a line, right click while the line is highlighted and select “Format Data Series”

The resulting graph should look like this

It’s a good start but some changes (i.e., adding a title, changing the lines, etc.) still need to be made.
Add a descriptive title
The title should contain information about both IVs and the DV

Selecting the bottom arrow will provide a drop down menu. Select the icon that shows crossed lines, a title, labeled axes, and a key. In this case, that is the tenth graph preview image. Doing this will add a space for a title and both axis.

These lines will need to be removed. This will be discussed on the next page.

With the graph highlighted, select the bottom arrow under “Chart Layouts”

Scores on Focus and Concentration Test as a Function of Amount and Time of Caffeine Consumption
Label the Y-Axis

To remove these lines, select the line and right-click

Select “Delete”
To remove the grid lines, select one line and right click

Select “Delete”
Reference