

Excerpt from
LABORATORY MANUAL
**PRINCIPLES OF PSYCHOLOGY:
EXPERIMENTAL FOUNDATIONS**

**PSYCHOLOGY 122
2001**

Participating Faculty

Professor James Dickson (dickson@stolaf.edu)
Professor Dana Gross (grossd@stolaf.edu)
Professor Elizabeth Hutchins (hutchine@stolaf.edu)
Professor Bonnie Sherman (sherman@stolaf.edu)
Professor Howard Thorsheim (thorshm@stolaf.edu)

Preceptors

Kristina Anderson
Katherine Audette
Boback Kristen
Gregory Amy

Experiment 6

Statistics Follow-up on Play Lab

Dana Gross

begins on the next page and constitutes pp. 74-89 of the full manual

Note: Author contact information: Dr. Bonnie Sherman, Psychology Department, St. Olaf College, 1520 St. Olaf Avenue, Northfield, Minnesota 55057 USA. (507) 646-2222 or 646-3146. Email: sherman@stolaf.edu

Copyright © 2002 by Bonnie Sherman, James Dickson, Dana Gross, Elizabeth Hutchins, & Howard Thorsheim. Used under license by the Society for the Teaching of Psychology (APA Division 2). All rights reserved. You may download and print 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 authors' names and institution and the Office of Teaching Resources in Psychology heading or other identifying information appear on the copied document. No other permission is granted to you to print, copy, reproduce or distribute additional copies of this material. Anyone who wishes to produce copies for other purposes must obtain the permission of the authors.

Statistics Follow-up on Play Lab

Dana Gross

There are three kinds of lies: lies, damned lies, and statistics.

--Disraeli

Statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write.

--H. G. Wells

Introduction

In your group's study of children's play, you used event sampling to record the occurrence of a particular behavior. The results of the event sampling tell you the total number of behaviors that occurred. While you were collecting your data, you may have noticed differences between your two groups. In this lab you will learn about descriptive statistics -- techniques for examining patterns and relationships in your data, and examining and comparing distributions. You will also discover how the variation of your data, the number of subjects you have, and the presence of outliers can affect these patterns and distributions. Finally, your group will plot your findings graphically and present your results as a poster.

StatView (the computer program you will use) enables you to determine (a) an overall arithmetic mean for the play behavior you focused on and (b) the means for each of your two groups (younger vs. older children or boys vs. girls). It will also give you an indication of the variability within each group.

Objectives

- To learn about the arithmetic mean and the indications of variability associated with it
- To analyze the data you have collected in your observational study of children's play behavior
- To use StatView to create a data file and compute descriptive statistics
- To use StatView to learn about techniques for graphically plotting data
- To develop a strategy for sharing the results of your observational study of children's play in a poster

Terms

Box plot

Cell bar plot

Data file

Descriptive statistics

Histogram

Mean
Median
Mode
Normal distribution
Outlier
Quartile
Range
Scattergram
Standard deviation
Variability

Methods

Materials

Data sheet used to record observations (1 per research team)
StatView statistics program
Computer disk (1 per research team)
Statistics Video “Why Use Statistics? Describing Data”

Note: It is important to bring your own disk to save your results from this lab

PART 1

Statistics Video, “Why Use Statistics?: Describing Data”

This video should solidify the concepts of the arithmetic mean, standard deviation, and a normal distribution that were introduced in lecture and covered in Peter Gray’s *Psychology*, A1-A9.

<i>CT</i> The data from your Play Project are quantitative. Suggest a way to gather qualitative data for the Play Project, using the same general topic you chose.
--

PART 2

Descriptive Statistics Using Play Project Data

Procedure

Opening StatView

1. The StatView program is located in the Psychology space in [Sven](#). (To get to [Sven](#), select the Chooser from the Apple menu. Select Software Servers and

choose **Sven**. You will need to type in your email user name and password. After doing this, find the **Sven**.MAIN icon on your computer's desktop and open it up. Within the Shared Spaces folder, open the Psychology folder.)

2. Within the Psychology folder, you can reach StatView by opening the following folders: (1) Applications, and (2) Statistics.
3. After you open the Statistics folder, launch the StatView program by double-clicking on the multi-colored StatView alias icon. (Be patient; it takes a while to load.)

Creating a Data File

Data files in StatView look like Charts A and B, which are shown below. The first sample data file shows how data for a study of age differences in pretend play would be organized. The second sample shows how data for a study of gender differences in pretend play would be organized. In both sample data files, the independent variable (either Age or Gender) is recorded in one column and the dependent variable (pretend play) is recorded in the other column. If you follow the steps below, the data file your group creates today should look like these data files.

	Age	Pretend Play	Input Column
1	Younger	4.000	
2	Younger	4.000	
3	Younger	5.000	
4	Younger	7.000	
5	Younger	7.000	
6	Younger	9.000	
7	Younger	8.000	
8	Younger	8.000	
9	Younger	9.000	
10	Younger	7.000	
11	Older	3.000	
12	Older	2.000	
13	Older	1.000	
14	Older	4.000	
15	Older	2.000	
16	Older	2.000	
17	Older	2.000	
18	Older	5.000	
19	Older	0.000	
20	Older	3.000	

Chart A

	Gender	Petend Play	Input Column
1	Boys	8.000	
2	Boys	9.000	
3	Boys	4.000	
4	Boys	7.000	
5	Boys	8.000	
6	Boys	8.000	
7	Boys	4.000	
8	Boys	6.000	
9	Boys	6.000	
10	Boys	8.000	
11	Girls	3.000	
12	Girls	4.000	
13	Girls	3.000	
14	Girls	4.000	
15	Girls	4.000	
16	Girls	4.000	
17	Girls	2.000	
18	Girls	1.000	
19	Girls	0.000	
20	Girls	5.000	

Chart B

- After opening the StatView program, choose the New option from the File menu. Untitled Dataset #1 appears on your screen, providing you with one column in which to enter your data.
- Move the cursor arrow to the words Input Column, click, and then type the name of your group (Age OR Gender).
- The title you have just typed appears at the top of the column. A new input column appears to the right of the column you have just labeled.
- In the second input column, type the name of the play behavior your group observed. (Put the arrow on Input Column, click, and type.)
- It is now time to enter your data. (Note: An empty input column appears to the right of your last data column. Disregard this empty column.)

Preparing for Data Entry

4. Before entering your data, you need to specify the type of data being entered. The default setting is real, which is appropriate for your second data column (i.e., the play behavior you observed). The first data column, however, needs to be changed to category data.
5. To do this, move the cursor to the box where the word real appears.
6. Hold down the mouse button and move the cursor arrow to the word category when it appears.
7. When you do this, a new box called Choose Category appears.
8. Choose the New button to label the first data column category.
9. An Edit Category dialog box will appear. In the box titled Group Label, type in the name of one half of your group. (For example, if your grouping variable is Age, type in “Younger.”) Then choose the Add button or just hit return.

10. Then type in the name of the other half of your group. (In this example, type in "Older.") Then choose the Add button (or just hit return) again.
11. Now select the Done button and StatView returns you to the data file.

Entering Your Data

4. You are now ready to enter the data you collected. Refer to the sample data file in this handout.
5. At the lower edge of your dataset, you will see a shaded row. This is where you enter subject data. More shaded rows will appear as you enter data.
6. For each child you observed, you need to enter the data you obtained on his/her (a) group category (Age OR Gender) and (b) number of target play behaviors he/she exhibited. (If the child did not exhibit the behavior, you should enter 0.) Click the shaded cell in the column (e.g., the Age column) where you wish to enter data. Type the category (e.g., "y" for "younger").
7. When you have entered the data for all the children in your sample, you are ready to perform the analysis of your data.
8. Save your data file and the results you obtain on your disk.

Descriptive Statistics

Describing Your Data Numerically

Descriptive statistics are a numerical summary of the data in your data file.

Descriptive statistics fall into three categories:

- measures of central tendency, which give an idea of the average value of a number or other quantity (e.g., mean, median, mode)
- measures of variability, which convey whether most measurements are clustered within a narrow range of values or spread over a large range (e.g., minimum, maximum, range, variance, standard deviation)
- measures of an overall distribution property indicated by a single number (e.g., skewness, kurtosis)

For the play project, you will focus on the first two of these categories -- measures of central tendency and measures of variability.

Example of descriptive statistics for Age

The descriptive statistics show that the overall mean (combining younger and older children) for pretend play was 4.60. The standard deviation and standard error are measures of variability around the mean. The count is the number of children in the data file, in this case 20 children. The minimum (0) and the maximum (9) show the range of pretend play behaviors observed. The # missing (0) shows that there are no missing data in the data file.

Descriptive Statistics

	Pretend Play
Mean	4.600
Std. Dev.	2.780
Std. Error	.622
Count	20
Minimum	0.000
Maximum	9.000
# Missing	0

Example of descriptive statistics for Gender

The descriptive statistics show that the overall mean (combining boys and girls) for pretend play was 4.90. The standard deviation and standard error are measures of variability around the mean. The count is the number of children in the data file. The minimum (0) and maximum (9) show the range of pretend play behaviors observed. The # missing (0) shows that there are no missing data in the data file.

Descriptive Statistics

	Pretend Play
Mean	4.900
Std. Dev.	2.532
Std. Error	.566
Count	20
Minimum	0.000
Maximum	9.000
# Missing	0

1. To compute descriptive statistics for your entire sample, select the Analyze menu and choose the option called Descriptive Statistics.
2. A box appears asking you to select a variable to be described.
3. In the right-hand box, your variables are listed. Choose (double click) the dependent variable, and then click OK or just hit return.
4. StatView creates a new Untitled View containing your descriptive statistics.
5. Print a copy of this view for each member of your group and save it on your disk.

Fill in the blanks in the tables below with information from your group's data.

Descriptive Statistics

	Play Behavior
Mean	
Std. Dev.	
Std. Error	
Count	
Minimum	
Maximum	
# Missing	

***CT* Observe the descriptive statistics for your study. Are you surprised by what you see, or does it confirm your expectations?**

PART 3
Plots of Data

Describing Your Data Graphically

Plots provide an effective way of looking at data to complement a numerical analysis. They allow you to see patterns and relationships among variables, spot outliers (extreme scores), examine and compare distributions of groups, and help you check assumptions about your data. Plots also offer a graphic way of presenting your analysis results, which aids you in communicating information about your data to others. Print out copies of your graphs and save them on your disc.

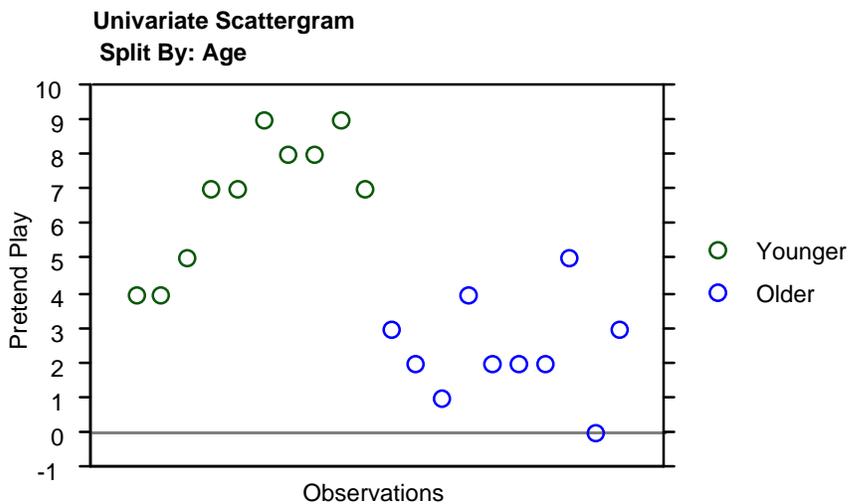
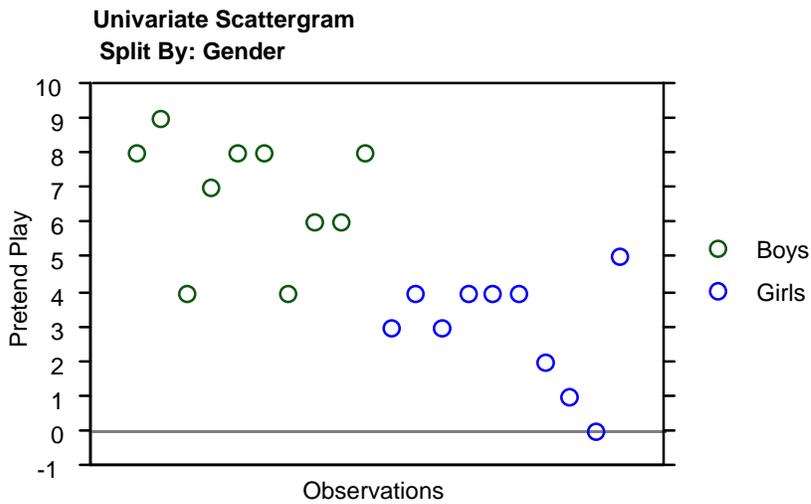
Graph 1: Univariate Plot

Univariate plots show the distribution of a variable in a one-dimensional plot with a single numeric axis, the Y-axis. Each observation is plotted along the horizontal axis in the sequence the data appears in the dataset. In StatView, you can display the observations as points in a scattergram, as points connected by lines in a line chart, or as bars in a bar chart. You can also distinguish different groups (e.g., boys vs. girls) in the plot.

To create a univariate plot of your data:

1. Select Graphs from the Analyze menu; then choose the option called Univariate Scattergram.
2. When the Univariate Plot box appears, select the dependent variable (the play behavior you studied).
3. Split the scattergram by your grouping variable by selecting the independent variable (either Gender or Age). Click OK. The two different groups should be

distinguished by different plotting symbols on the scattergram. Your graph should look like the examples below.



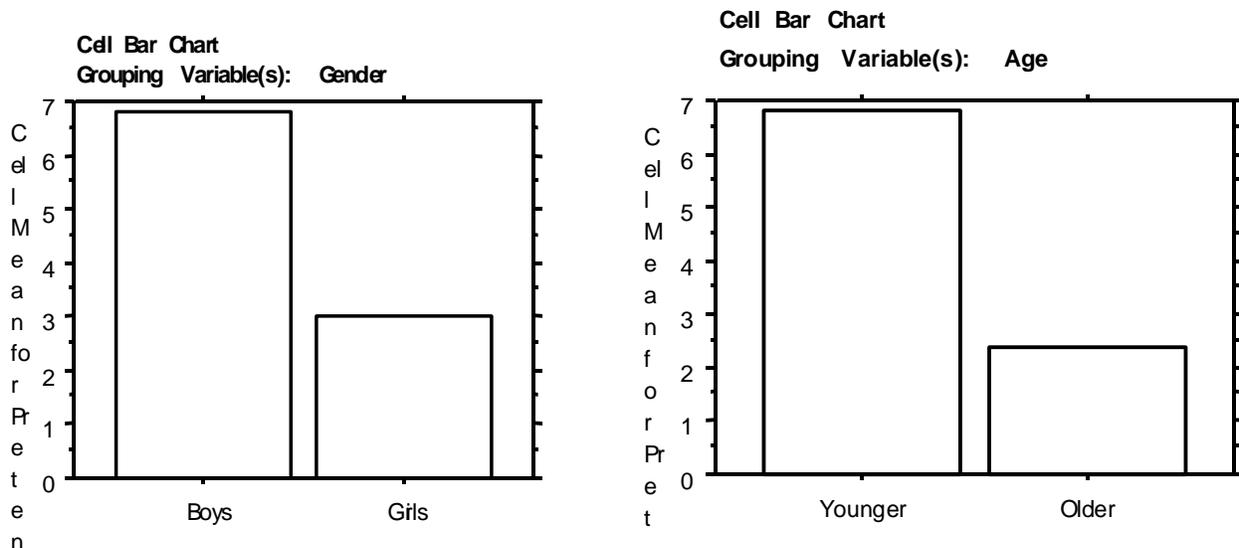
CT Do the distributions for each of your groups seem to be similar or different from each other?

Graph 2: Cell Bar Chart

An easy way to plot and compare data for two groups (e.g., boys vs. girls) is to construct a bar chart. In a bar chart each group is represented by a bar with a height equal in magnitude to some measure of that variable. One of the most effective measurements in event sampling is the mean.

To construct a bar chart for your data make sure your view window for analyzing data is open. This is the same window in which you performed your descriptive statistics.

1. In the left-hand column, scroll down until you see Cell Plots. Click on the arrow to bring down the menu.
2. Double click on the Bar Chart option.
3. When the Cell Plot dialog box appears, click OK.
4. In your variable browser box, select your grouping or nominal variable (Age or Gender) and then click Add.
5. Then, in your variable browser box, select your play behavior or continuous variable and then click Add.
6. Your bar graph should then be complete. It should look similar to the examples below.



CT Does the cell bar chart for your data seem to support your interpretation of the univariate plot?

Graph 3: Histogram

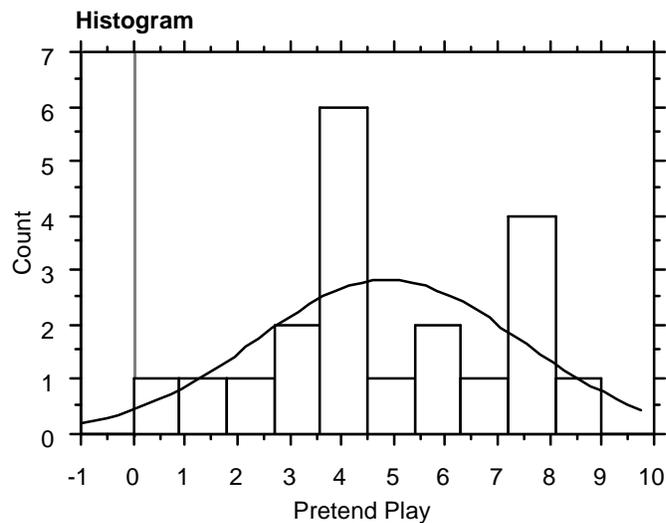
Histograms are a way of showing the distribution of data in comparison to the normal curve. The abscissa is divided into multiple numerical classes of increasing magnitude. Over each class, a bar appears equal in height to the frequency of that class. Histograms are an effective way to notice outliers.

To construct a histogram for your data, make sure your view window for analyzing data is open. This is the same window you used to construct your bar graph.

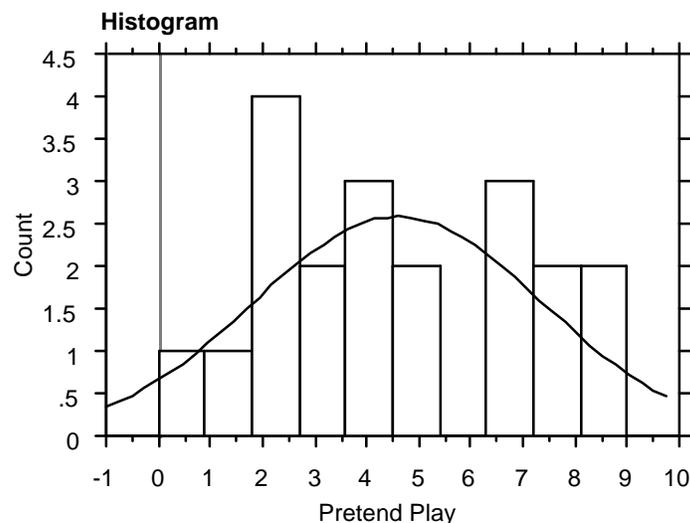
1. In the left-hand column, scroll down to Frequency Distribution and click on the arrow to see the drop down menu.

2. Double click on Histogram.
3. When the Frequency Distribution dialog box appears, make the following changes:
 - a) in the box labeled Number of Intervals, enter the maximum number of events any child had, regardless of which of your two groups they belonged to.
 - b) Click on the gray box labeled Show Normal Comparison. A check should appear.
 - c) Click OK.
4. You should see an incomplete box on your screen.
5. In your variable browser box, select your continuous (play behavior) variable and then click Add.
6. Your graph should appear, along with the normal curve. It should be similar to the examples below.

Age:



Gender:



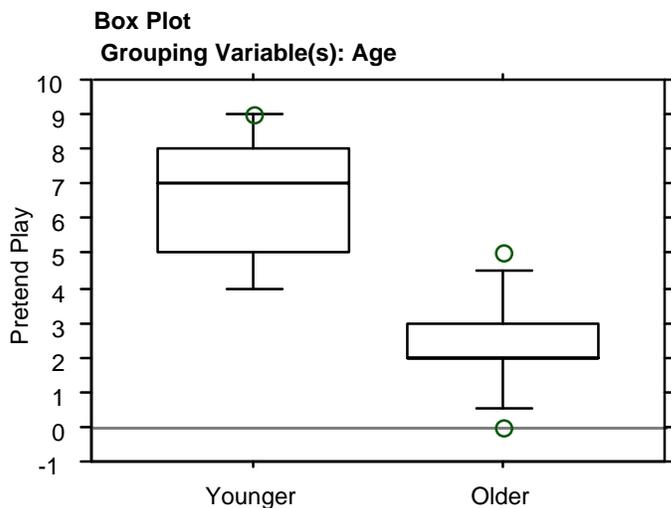
CT Based on the histogram for your data, does the play behavior you observed appear to be normally distributed?

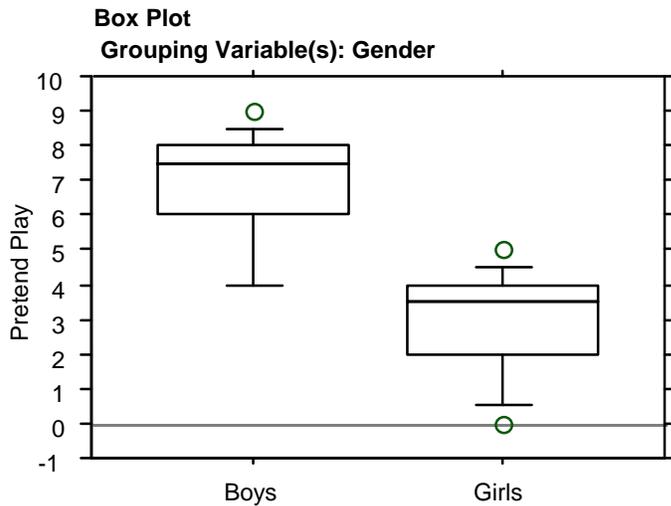
Graph 4: Box Plot

Data that contain outliers may be measured more effectively by the median and its respective tests, than by the mean. A box plot measures the distribution of the data based on the median rather than the mean. The bar within the box represents the median while the edges of the box represent the quartiles. The lines outside of the box, at the edge of the figure, measure the spread.

To construct your own box plot:

1. From the Analyze Menu, choose Graphs, then Box Plot.
2. In your variable browser box, select your continuous (play behavior) variable.
3. Next select your nominal (grouping) variable for the axis variable. Click OK.
4. Your graph should now have two box plots on it, one for each of your groups. It should be similar to the examples below.





Part 4 The Effects of Fluctuations in Data

Statistical results are greatly affected by several characteristics of the data. These characteristics include: 1) sample size 2) standard deviation and 3) the presence of data points that differ drastically from the mean (outliers). We will manipulate these three characteristics to test their effect on your results. **NOTE: This is simply an investigation of the mechanics behind data distributions. It is never acceptable to alter scientifically derived data in order to obtain significance.**

Record the necessary data and answers in your lab notebook. Print out copies of your graphs and results and save them on your disc.

Experiment 1: Altering the sample size

Start a new table of data. Pretend every child you studied is one of three children who all behaved in exactly the same manner. Therefore, for every one data point you originally had, you now have three. For example if you had the original data points:

1,4,3,2,5,4,3,4,5,3,2 (so the first child had 1 event, the second child had 4, etc.),
your new data points would be:

1,1,1,4,4,4,3,3,3,2,2,2,5,5,5,4,4,4,3,3,3,4,4,4,5,5,5,3,3,3,2,2,2 (so now the first **three** children had 1 event, the next **three** had 4 events, etc.)

Using these new, altered data, repeat the descriptive statistics and the plots you created with your actual data.

**CT What do you notice about the mean and the variability?
Are they the same or different as before?**

CT How do the graphical representations compare to those derived from your original data?

Experiment 2: Increasing the standard deviation

As shown in the statistics video, standard deviation measures how far values lie from the mean in a distribution. It is calculated using the following equation:

$$S = \sqrt{1/(n-1) \sum (x_i - \bar{x})^2}$$

For each subject in your study, square the number of events recorded. For example, if subject one had three events originally, he/she would now have nine events (3^2).

Do this for all subjects and run your descriptive statistics and plots again.

CT What was the effect on your descriptive statistics?

CT How do these graphical representations compare to those derived from your original data?

Experiment 3: Outlier Exercise

1. Identify which of your two groups had the larger mean. Add 3 subjects to that group. These three “new” subjects should each have zero behavioral events.

Repeat the descriptive statistics and the plots with this new data set.

CT What effect did this have on your data?

CT How do these graphical representations compare to those derived from your original data?

2. Instead of adding three new subjects with values of zero, now add one subject who has 40 events.

Repeat your descriptive statistics and the plots again

CT What effect did this value have on your data?

CT How do these graphical representations compare to those derived from your original data?

Results

In APA-style research reports, results of statistical analyses appear in the Results section. Your poster should include numerical descriptions of your findings: the arithmetic mean, standard deviation, and the range of the data. Your poster should also include two graphic descriptions of your results: the cell bar chart and one of the other graphs (your group can choose the second graph of its choice).

Sharing the results

When sharing the results of investigative research, the outcomes of statistical analyses are always presented within a context. As a result of reading journal articles throughout the semester, you know that psychologists follow conventions of APA style when presenting the context for their findings. Information in APA-style publications is organized into a number of distinct sections: abstract, introduction, method, results, discussion, and references. These conventions are also used in posters reporting the results of empirical studies. Your group's objective during the remainder of this lab is to determine how you will present the results of your study in a poster.

Six Basic Steps for an Effective Poster Presentation

1. Define your audience

To ensure that your poster communicates as effectively as possible, it is important to understand who your audience is. Consider that the audience

- may not be as familiar with your topic as you are
- must be drawn in with your key findings
- has only a short time to look at your poster

2. Gather the content for your poster

Divide the information you have gathered into the following sections:

- Title: name of project/poster, contributors, name of class, college
- Introduction: description of the topic, relevant previous research (cite the author(s) and year of publication), hypothesis, operational definition
- Method: **brief** description of the process (use separate subheadings for: participants, materials, and procedure) Make sure that your poster has a "Method" label for this section.

- Results: summary descriptive statistics and plots of your data (Do not include the raw data file from your study.) For each group you studied, your poster should include the arithmetic mean, standard deviation, and the range of the data. Include your cell bar chart and one of the other graphs (your group is free to choose which of the other graphs to include). You **should not** interpret your results or explain what they imply—just state what they are. Make sure that your poster has a “Results” label for this section. In addition, graphs and figures need to be labeled; be sure to label both axes on all graphs.
- Discussion: interpretation and discussion of key results, a few easily remembered conclusions, topics for additional research, possible sources of error. Relate your findings to the research described in the Introduction by making explicit comparisons. Cite sources using APA style: author(s) and year of publication. Make sure that your poster has a “Discussion” label for this section.
- References: references for cited material in APA style. Make sure that your poster has a “References” label for this section.

Note: The posters will be graded by both preceptors from your section (A or B) and the four faculty members responsible for running the labs. In addition to the content, be thoughtful about the overall presentation -- neatness, readability, clarity, balance, and aesthetics!

3. Make a heading and statement for each section. Take each of the sections and summarize its content. Each section should be divided into three categories.

- Heading: title
- Statement: clear and succinct text relating to the heading
- Support material: documentation, photographs, maps, diagrams, charts, illustrations, etc.

4. Eliminate noise

After making the first draft of the poster, significant editing will probably be needed.

- Think about what will be most interesting to your fellow researchers.
- Pare down the detail, eliminating all but the vital elements.
- Reduce your information to brief, legible statements--**EVERYTHING** should be legible from 3 feet away!
- Use supporting material to encourage and facilitate understanding.

5. Find your focus

Because your audience has limited time to view your poster, ask yourself about the main points you would want the observer to come away with. There may be extra information you would like to have available; you can put together a sheet of more in-depth descriptions for those interested.

6. Put the poster together.

Discussion Questions

1. Based on the previous exercises, what three factors affect statistical outcomes? Which factor most prominently affected your data? How would you control for this if you were to repeat your study?
2. Which graphic representation do you find most informative? Explain why you prefer this one to the others.

References

Required Lab Reading

Paulos, J. A. (1990). Pseudoscience. In *Innumeracy: Mathematical illiteracy and its consequences* (pp. 49-71). New York: Vintage.

Suggested Readings

Gonick, L., Smith, W., and Smith, W. (1993). *The cartoon guide to statistics*.

New York: HarperCollins. [These inspired cartoons simplify the confusing concepts of statistics.]

Huff, D., & Geis, I. (1993). *How to lie with statistics*. New York: Norton. [This 1954 classic, which has been reissued, informs the statistically naive about the way advertisers, government, and the media can mislead their audiences through the misuse of statistics.]

Paulos, J. A. (1990). *Innumeracy: Mathematical illiteracy and its consequences*. New York: Vintage. [Paulos examines many aspects of popular culture, from stock scams and newspaper psychics to diet and medical claims to demonstrate the popular misconceptions resulting from the inability to deal with large numbers, probability, and ratios.]

Paulos, J. A. (1996). *A mathematician reads the newspaper*. New York: Anchor. [This humorous book reveals the hidden mathematical angles in media stories.]

<http://www.stats.org/> Statistical Assessment Service. [The Statistical Assessment Service (STATS) is a non-partisan, non-profit research organization in Washington, D.C. STATS is devoted to the accurate use of scientific and social research in public policy debate. STATS serves as a resource for journalists by providing timely and well-researched analysis of current statistical and scientific disputes.]