Ψ120 Research Methods (Van Selst): Laboratory #1 (Graphing)

Cozby + APA Manual – This June 3 lab is due at the start of class on June 10, 2014

Bar-Charts, Histograms, Polygons, Relative Frequency Plots, & Plots of Cumulative Frequency.

Simple and grouped frequency distributions provide useful summaries of data. Often this information is easier to interpret when presented in graphic form rather than in table form.

Frequency distributions are often portrayed graphically in the form of *bar-graphs* or *histograms*. The class intervals (or scores) are shown on the horizontal axis (*the abscissa*). The frequencies of each class interval (or score) are shown on the Y-axis (*the ordinate*).[1] Consider the following data showing predicted coffee consumption at various ages for a hypothetical population.

Income (000's)	10-11,999	12-13,999	14-15,999	16-17,999	18-19,999	20-21,999	22-23,999	24-25,999
Count	2	2	4	4	5	1	2	1

Figure 1.1 illustrates the data with bars that do not touch (i.e., a *bar-chart*). This implies that the categories are qualitatively different (as opposed to quantitatively different). Thus, a bar-chart (e.g., Figure 1.1) is probably inappropriate for this data. Figure 1.2, a histogram, is better. Bars that touch (i.e., a Histogram) imply that the categorical variable on the X-axis (the abscissa) uses an interval or ratio scale (i.e., a continuous underlying variable with scale values representing equal intervals).

Figure 1.1

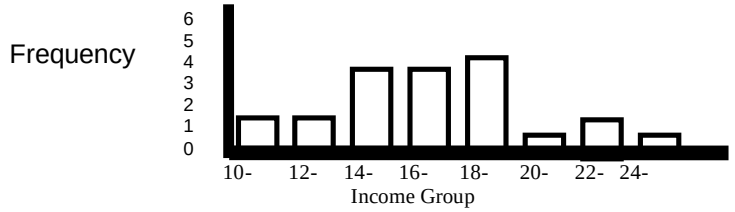
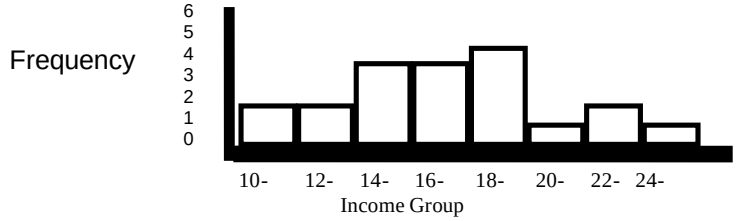


Figure 1.2



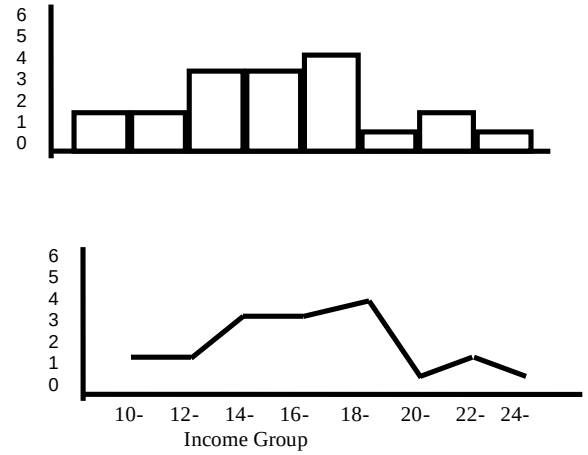
The visible difference between the bar-graph (Figure 1.1) and the histogram (Figure 1.2) is whether or not the bars touch. The theoretical importance of this difference is that whether or not the bars touch reflects whether the variable plotted on

[1] Note: there are a number of heuristics to remember which the abscissa is and which the ordinate is. One is that X comes before Y and Abscissa ("a") comes before Ordinate ("o"). A second is to use your mouth position after the first syllabus: "AB" → closed mouth → X-axis and "OR" → open mouth → Y-axis.

the abscissa[2] is continuous (Figure 1.2, histogram) or not (Figure 1.1, bar chart). [in this case the histogram is the correct alternative]

Often frequency plots are shown as frequency polygons (histogram) or line-graphs (Figure 1.3). The axes are labeled just as they would be for a histogram but the data points are used as anchor points for lines rather than to indicate the height of the bars for the histogram (compare Figure 1.3 vs. Figure 1.2).

Figure 1.3



Other notes:

- If you do not show the origin (the zero-point) on the ordinate it is important that you indicate this by "breaking" the axis right before the intersection point (often with a double-slash or pair of "lightning bolts").

- There is a subtle issue of having a graph represent the information clearly and having the graph misrepresent the data. Consider the following graphs. The upper panel shows crime rate drops sharply, the lower panel shows only a modest decrease, yet both graphs illustrate the same data. Although both of the following graphs are "accurate", it should be clear that they do not pull for the same conclusions. The "best" graph of this data somewhere between these two extremes.
- Ensure that the zero values are included where they exist in the data.

[2] For graphs with a continuous underlying variable on the abscissa (histograms, line-graphs) aim for a breakdown of eight to twelve intervals (unless there is a good theoretical or methodological reason not to).

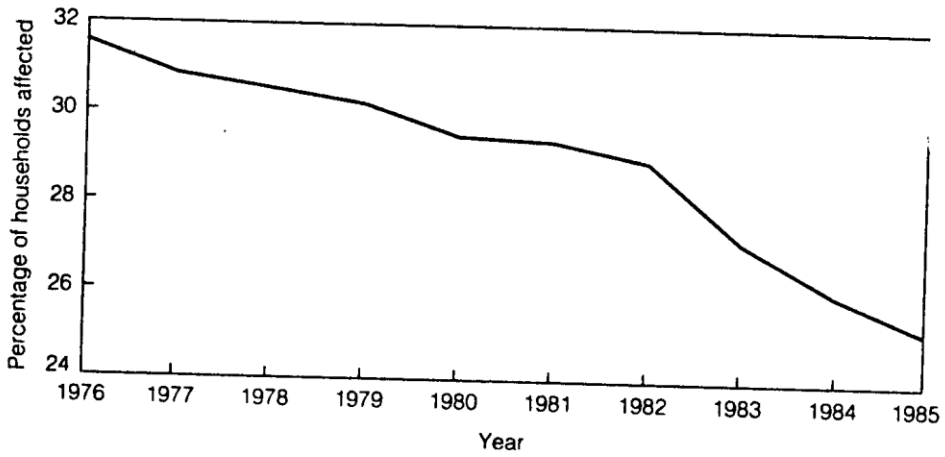


Figure 4.3 Crime rate drops sharply

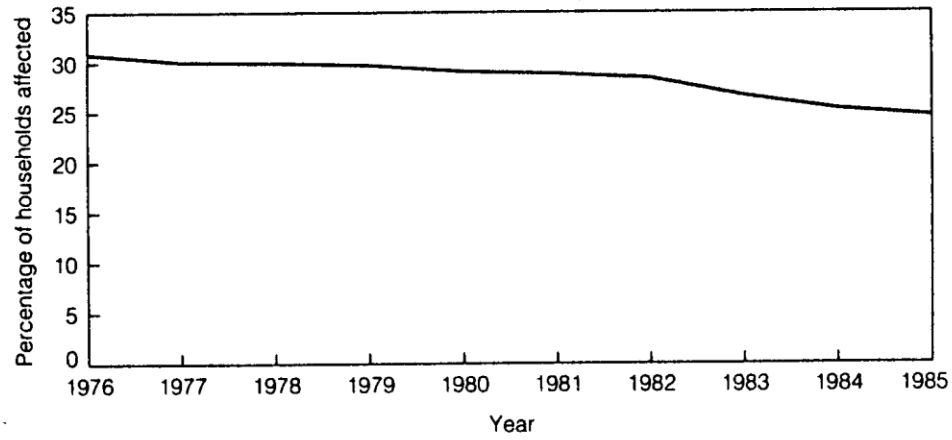


Figure 4.4 Crime rate drops slightly

Instructions for Data-Collection

This assignment will be graded based on the data collected in the laboratory section.

You will do your data collection in partner-pairs (this also provides you with an opportunity to meet your classmates).

There are three tables on the following sheets. Each table has eight rows (subject number & seven variables). For now, just concern yourself with Table 1. We will fill-in the big table as a class prior to the end of the laboratory session.

Table 1.

Subject #	Memory Test #1	Memory Test #2	Memory Game	Luck #1 (1d6)	Luck #2 (# rolls)	Shoe Size	Height (inches)
YOUR DATA	SC= LC=	SA= LC=					

Variable #1: MEMORY TEST #1

The instructor will read a list of words to you. Your task will be to memorize (and then write down after a two minute delay) as many of these words as you can. Thus:

1. put down your pen
2. listen carefully during the presentation of the words
3. QUIETLY wait for the Instructor to tell you to start writing
4. Write down as many of the words as you can remember.
5. after two minutes (time enough to recall as many words as you can), you will note the number of words you correctly recalled
6. to score your data:
    - put the number of words correctly recalled in the MEMORY TEST #1 box. (Note: consider plural forms or other variants of the base word as "correct").
    - put words which you recalled but which were not in the original list [i.e., inclusion errors] in brackets
    - For example, if you recalled six words (and one additional one not from the set), you would write 6(1) under MEMORY TEST #1 for "my own data".

Variable #2: MEMORY TEST #2

Same as #1 but with a different set of words.

Variable #3: MEMORY GAME
Your instructor will provide some six-sided dice. Read over these instructions then:

1. Take a single six-sided die. Roll it once. This is your first number.
2. Read it aloud to your partner (your partner will write these numbers down on a sheet of paper hidden from your view).
3. Roll the die again. Recite the first and second ("new") number to your partner.
4. Continue rolling & reciting numbers until your partner tells you that you have made a mistake. Note: You must guess if you cannot remember.
5. Write down the number of items that you got "correct". For instance, reading the first roll (reading the number) gives you a score of "1". The second roll (recalling roll#1, reading the new number) will get you a score of "two". If you mess up on the third roll (e.g., cannot remember the second roll), this means that you write down a score of "2" on "my own data" for the MEMORY GAME variable.

Variable #4: Luck #1
Roll a six-sided die. Record the value of the number that appears on the die as your Luck #1 score (note: there are no "practice rolls" or "do-overs", record the first number that you roll).

Variable #5: Luck #2
Roll a six-sided die. If the number showing on the die is a five or six, stop. If not, continue to roll the die until you EITHER roll a five or six OR you roll the same value as the immediately preceding roll. ***Your "Luck #2" score is the total number of rolls (including the last roll)***. It is NOT the total of the values.

- Example 1: roll a five. Stop. LUCK #2 = 1.
- Example 2: roll a two, one, four, five. Stop. LUCK #2 = 4
- Example 3: roll a one, two, one, five. Stop. LUCK #2 = 4
- Example 3: roll a one, one. Stop. LUCK #2 = 2.

(note: there are no "practice rolls" or "do-overs", record the first LUCK #2 score that you obtain).

Variable #6: Shoe Size.
Write down your (American) shoe size.

Variable #7: Height.
Write down your height (in inches).

Table 2.
Lab-section data-sheet.

Participant Number	Memory				Memory Game	Luck #1 (1d6)	Luck #2 (# rolls)	Shoe Size	Height (inches)
	SC	LC	SA	LA					
1									
2									
3									
4									
5									
6									
7									
8									
9									
10									
11									
12									
13									
14									
15									
16									
17									
18									
19									
20									
21									
22									
23									
24									
25									
26									
27									
28									
29									
30									

Ψ120 Research Methods (Van Selst): Laboratory #1 (Graphing)

The Assignment

Use graph paper for all graphs in this assignment (and for any hand-drawn graph in any future assignment). **You must use one sheet of paper for each graph and more than 70% of the available space on each sheet.** Being both accurate and neat is important. You will need to use a ruler.

1. For the variable "MEMORY TEST #1 (SC, LC)", graph the distribution of the summed scores of SC + LC for each individual. You will need to compute the totals for the data from each participant individually. Once you have the full data set, build a frequency table and histogram[3] — *a graphic form of a frequency distribution in which the number of cases within each class is represented by a vertical bar whose height is proportional to the number of cases within a class* — (for a histogram the bars should touch: see Fig. 1.2 of this hand-out, page 244-245 in the Cozby text). Describe the resulting graph in words, including the minimum and maximum scores, the range of scores, and the mean, the median, and mode for the group on that variable (Cozby p. 245-246). Note: for this assignment, do not bin the scores for your histogram. Also, you MUST include the frequency table (e.g., two columns: one labeled score, the other frequency) as well as the figure. The scores should be in ascending order. The frequency should indicate the number of individuals who obtained each score. It is ok, but not required, to have the frequency table on the same page as your histogram.

e.g.,

Score	Frequency
3	0
4	2
5	2
6	1
7	3
…	…

2. Make a graph of the (overall) **mean** performance of SC vs the (overall) mean performance of LC. Use "Word Length" as the label on the abscissa (with headings of "short (4)" and "long (5-7)" for SC and LC respectively) and **NUMBER OF WORDS RECALLED CORRECTLY (/12)** (i.e.,

[3] *Each graph should include a title, axis-labels for each of the abscissa (X-axis) and ordinate (Y-axis), the value-labels on each axis, a mark to indicate the precise location of the value (e.g., the "dash" beside each time interval).*

"hits") on the ordinate. Figure 1.1 of this hand out provides a model to base your work on (note: separate bars for the bar chart). Be sure to indicate the title, etc. as before. Also, be sure to indicate the unit of measurement.

3. Produce a scatterplot to illustrate the relationship between the two measurements of body length: shoe size and height (inches). For scatterplots use the raw scores. Do not collapse into larger intervals (each subject produces only a single data point on the scatterplot).
4. Produce a scatterplot to illustrate the relationship between LUCK #1 and MEMORY GAME. For this graph each student will contribute 1 data point. You may have to develop a representational scheme for multiple overlapping data points (e.g., a "•" for 1 observation, "(2)" for two overlapping observations, "(3)" for three overlapping observations, etc.).
5. For the two scatterplots (questions #3 & #4), describe the relationships that you have plotted in words. Which two variables appear to have a stronger relationship? Discuss why that may be.
6. Discuss a possible relationship between SC and LC. How might you assess whether this relationship exists?
7. Produce a scatterplot (see question 4 above) between SC and LC.
8. How do the observed distributions of the "LUCK" measures differ? Discuss how you might defend your answer.
9. Replicate the graph for question 2 using the graphing capabilities within either word or excel.
10. Replicate the graph for question 7 using the graphing capabilities within either word or excel.

- -

After data collection you will transfer your data to the sheet at the front of the room (the laboratory copy of Table 2). The instructor will make copies for everyone in the class.