

Basic Descriptive Statistics Healey, Chapter 2

Percentages, Ratios and Rates,
Frequency Distributions, Charts and
Graphs

Outline:

- Percentages and Proportions
 - Ratios, Rates, and % Change
 - Frequency Distributions
 - Charts and Graphs
-

Percentages and Proportions

Formulae:

$$\text{Proportion: } p = \frac{f}{N}$$

$$\text{Percentage: } \% = \left(\frac{f}{N} \right) \times 100$$

where f = frequency, or the number of cases in any category

N = the number of cases in all categories

Percentages and Proportions (cont.)

- Report ***relative*** size.
- Compare the number of cases in a specific category to the number of cases in *all* categories.
- Compare a *part* (*specific category*) to a *whole* (*all categories*).
 - The *part* is the numerator (f).
 - The *whole* is the denominator (N).

Percentages and Proportions (cont.)

- Suppose you have a group of 229 sociology majors, of which 97 are female and 132 are male.
- What percentage of this group is female?
 - The *whole* is the number of people in the group.
 - The *part* is the number of females.

Percentages and Proportions (cont.)

- To identify the whole and the part, use the keywords **of** and **is**.
 - **of** identifies the **whole** (N)
 - **is** identifies the **part** (f)

Percentages and Proportions: Example

- What % of social science majors is female?
 - of (whole) = all sociology majors
 - $97 + 132 = 229$
 - is (part) = female sociology majors
 - 97
 - $(97/229) * 100 = (.4236) * 100 = 42.36\%$
 - 42.36% of sociology majors are female

Ratio

- Compares the relative sizes of categories.
- Compares ***parts*** to ***parts***.
- $\text{Ratio} = f_1 / f_2$
 - f_1 - number of cases in first category
 - f_2 number of cases in second category

Ratio (cont.)

- In a class of 23 females and 19 males, the ratio of males to females is:
 - $19/23 = 0.83$
 - For every female, there are 0.83 males.
- In the same class, the ratio of females to males is:
 - $23/19 = 1.21$
 - For every male, there are 1.21 females.

For Practice:

- With a partner, try Healey, #2.1 in 2/3e (a-e)
- a. % married in A ____ ? in ____ ?
- b. Ratio single to married in A ____ ? in B ____ ?
- c. Proportion widowed in A ____ ? in B ____ ?
- d. % single living in B ____ ?
- e. Ratio unmarried (living together) to married in A ____ ? in B ____ ?

Rate (cont.)

- Expresses the number of actual occurrences of an event (births, deaths, homicides) vs. the number of possible occurrences per some unit of time.
-

Rate (cont.)

- Birth rate is the number of births divided by the population size times 1000 per year.
- If a town of 2300 had 17 births last year, the birth rate is:
 - $(17/2300) * 1000 = (.00739) * 1000 = 7.39$
 - The town had 7.39 births for every 1000 residents.

For Practice:

- With a partner, try Healey #2.3
 - Rate for bank robberies? _____
 - Rate for murders? _____
 - Rate for auto thefts? _____
-

Percentage Change

- Measures the relative increase or decrease in a variable over time.
- Formula:

$$\text{Percentage change} = \left(\frac{f_2 - f_1}{f_1} \right) \times 100$$

where f_1 = first score, frequency, or value

f_2 = second score, frequency, or value

Percentage Change (cont.)

- f_1 is the first (or earlier) frequency.
 - f_2 is the second (or later) frequency.
 - Change can also be calculated with proportions, rates, or other values.
-

Percentage Change: Example

- In 1990, a city had a murder rate of 7.3.
-
- By 2000, the rate had increased to 10.7.
- What was the relative change?
 - $(10.7 - 7.3 / 7.3) * 100 = (3.4 / 7.3) * 100 = 46.58\%$
- The rate increased by 46.58%.

Frequency Distribution

- This is a report in the form of a table of the number of times each score of a variable occurred.
 - The categories of the frequency distribution must be stated in a way that permits each case to be counted in one and only one category. Categories must not overlap (they should be “mutually exclusive.”)
 - Table should have a title and clearly labeled categories and columns.
-

Example: Frequency Distribution for Age

(Interval width = 2 years)

Class Interval Age	Frequency	%
18-19	11	55
20-21	5	25
22-23	2	10
24-25	1	5
26-27	1	5
Total	20	100

Stated Class Limits and Real Class Limits

- In the previous table, the limits of the intervals appear to have a “gap” between categories: the scores of the variable are organized into discrete intervals. These are the **stated class limits**.
- Some calculations require that the “gap” be eliminated so the intervals appear continuous. To do this, you have to find the **real class limits** by adding half the distance to the upper limit, and subtracting half the distance from the lower limit. In the **Age** table, the gap is equal to 1, so you would add and subtract half that distance, or .5, to either end of the interval.

Midpoints

- You will also need to find the midpoints of the intervals for some statistical calculations and for graphing purposes.
 - The midpoints can be found by adding the upper and lower limits together and then dividing the total by 2.
 - See **Age** example below for the real class limits and midpoints.
-

Example: Frequency Distribution for Age

Class Interval Stated Limits	Real Class Limits	Midpoints
18-19	17.5 - 19.5	18.5
20-21	19.5 - 21.5	20.5
22-23	21.5 - 23.5	22.5
24-25	23.5 - 25.5	24.5
26-27	25.5 - 27.5	26.5

Cumulative Frequency and Percentage

Class Interval Age	Frequency	Cumulative Frequency	%	Cumulative %
18-19	11	11	55	55
20-21	5	16	25	80
22-23	2	18	10	90
24-25	1	19	5	95
26-27	1	20	5	100
Total	20		100	

Grouping Data (Interval-Ratio)

- When you have interval-ratio data, you may have many scores to put into a frequency distribution, so the data may have to be grouped into intervals with widths of 5, 10, or sometimes more, depending on the range of the scores.
- All intervals should be equal in size and should not overlap.
- Do not use more than 15 intervals (10 intervals is a good “rule of thumb” to follow.)
- Once you have decided on your interval width and number of intervals, construct the table in the same way as you would for nominal and ordinal data.

For Practice:

- Try Healey, #2.5. Construct a frequency distribution for the variables **Sex and Age**, including a column for %.
 - Complete this question as part of your homework.
-

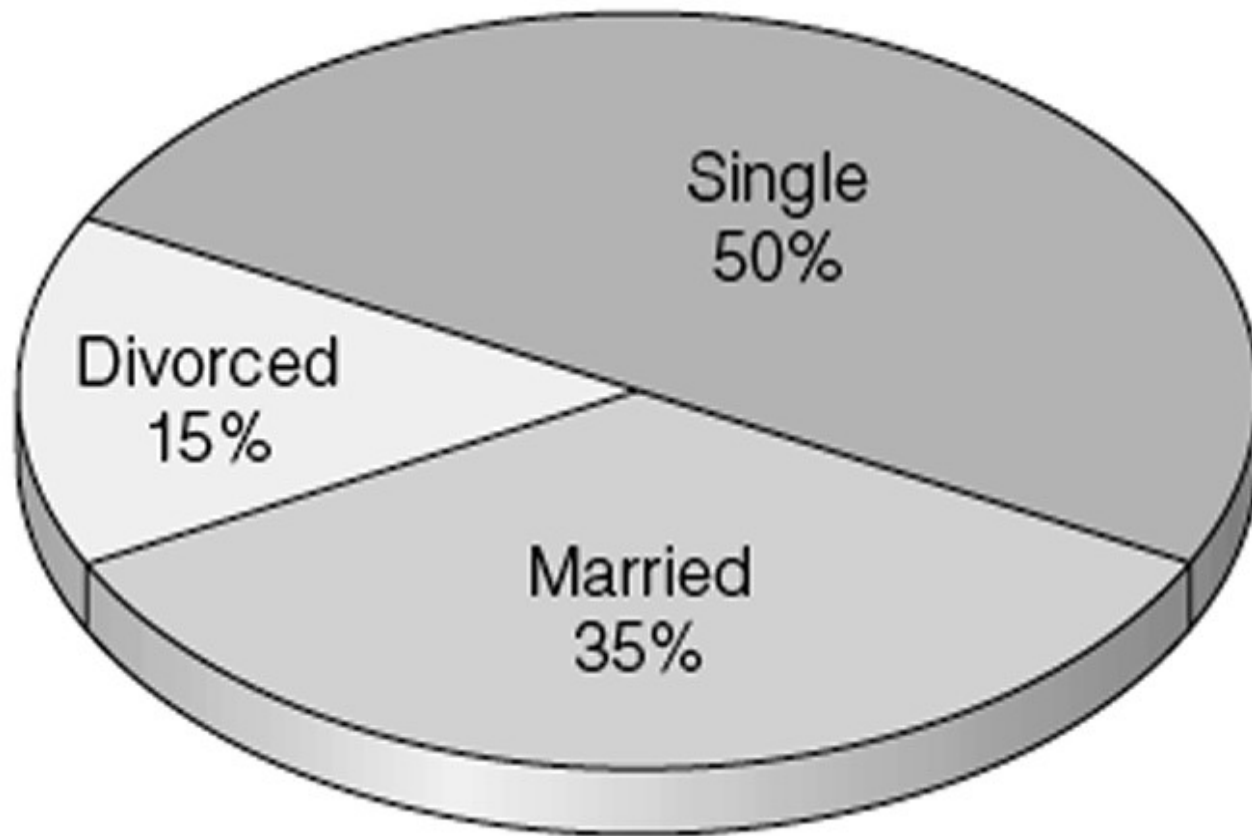
Graphs And Charts

- Histograms, pie and bar graphs and line charts (also called “frequency polygons”) can be constructed to present frequency distributions graphically.
 - Graphs and charts are commonly used ways of presenting “pictures” of research results.
 - Graphs can be constructed by hand or can easily be generated using a software program like Excel or SPSS.
-

Graphs and Charts (cont.)

- Histograms and frequency polygons are commonly used for interval-ratio data.
 - Bar graphs and pie charts are most often used for nominal or ordinal data.
 - Read rest of chapter in Healey for details on how to construct graphs and charts.
-

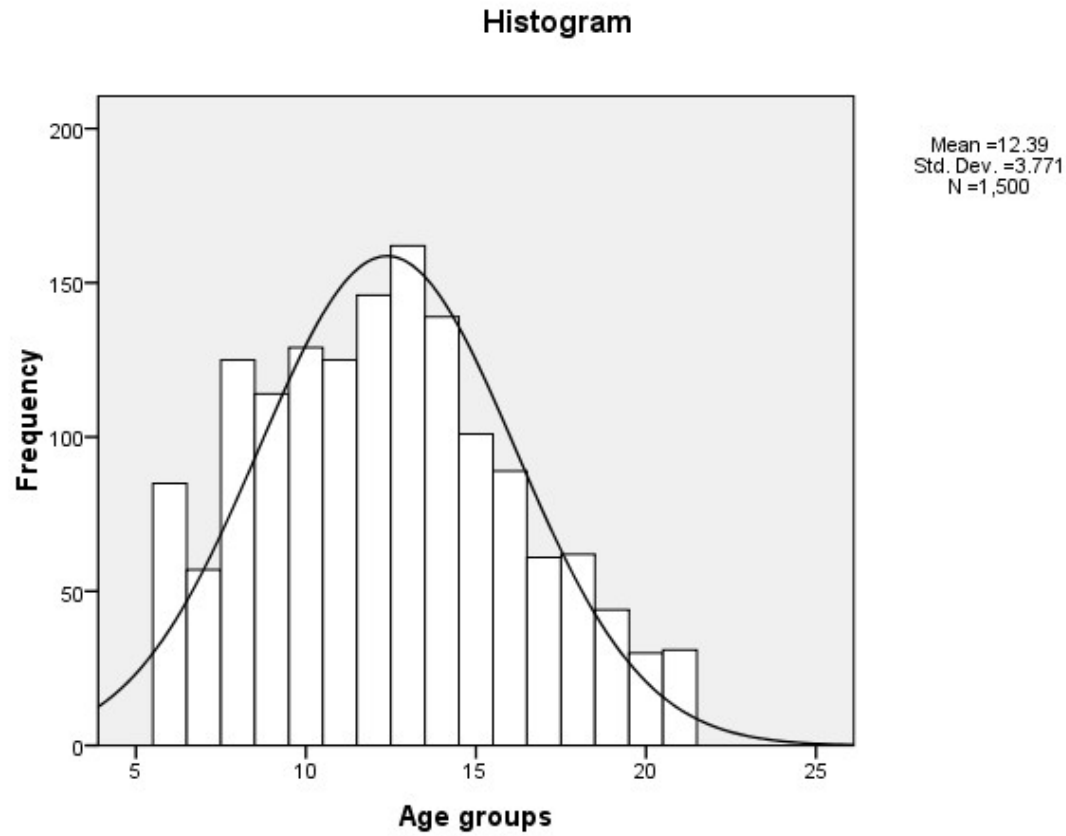
Sample Pie Chart: Marital Status (N = 20)



Sample Bar Chart: Marital Status Of Respondents (N = 20)



Histogram for Age Group



Marriage And Divorce Rates Over Time

How would you describe the patterns?



Homework Questions:

- 1. Complete Healey and Prus #2.5
 - 2. Healey, #2.2
 - 3. Healey, #2.9
 - SPSS Exercise:
 - Read SPSS section at end of Ch. 2
 - In the computer lab, try #2.1 and #2.2 (in SPSS section) for practice
-