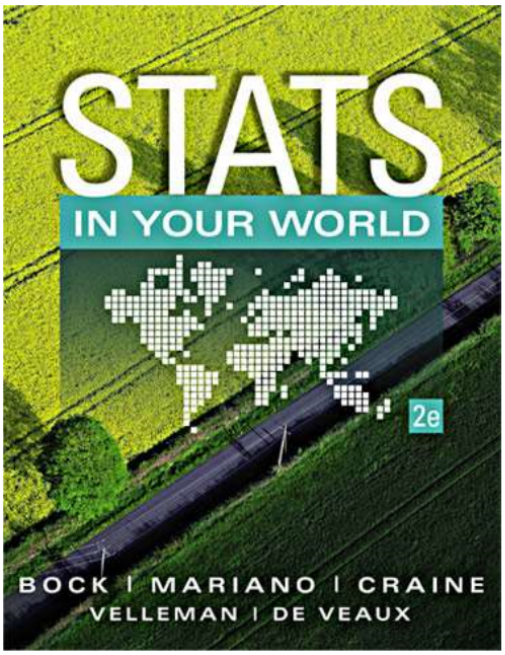
Chapter 3

Exploring Quantitative Data



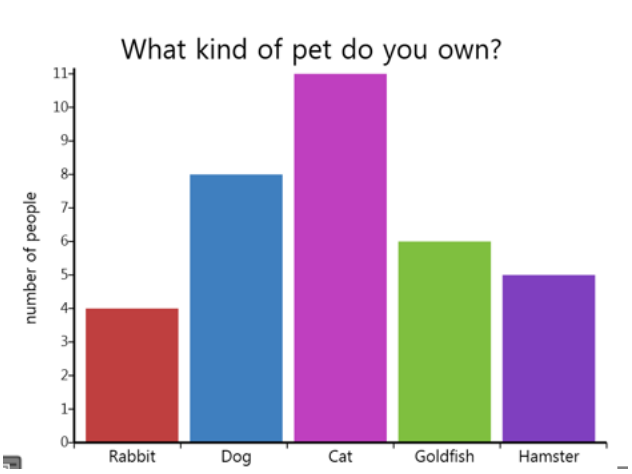
Day 1 Hwk: Histogram WS #1-10

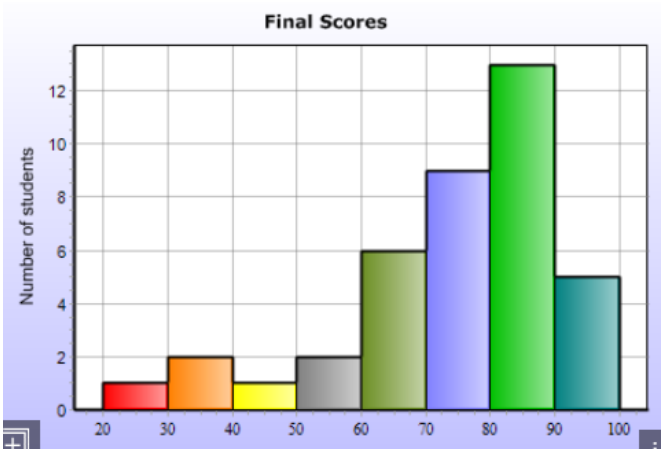
Day 2 Hwk: Histogram WS #11-18

Dealing With a Lot of Numbers…

- Summarizing the data will help us when we look at large sets of quantitative data.
- Without summaries of the data, it's hard to grasp what the data tell us.
- The best thing to do is to make a picture…
- We can't use bar charts or pie charts for quantitative data, since those displays are for categorical variables.

Bar Chart or Histogram?
Categorical or quantitative data?



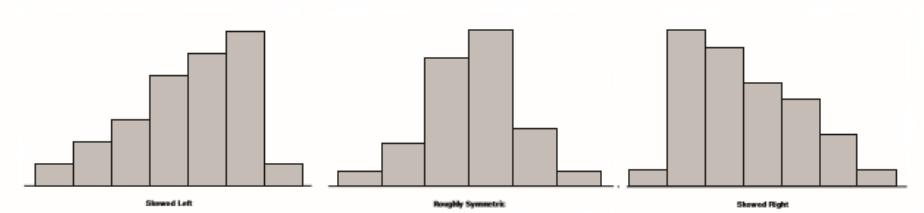


KEY TERMS

A **frequency distribution** provides a means of organizing and summarizing data by classifying data values into class intervals and recording the number of data that fall into each class interval.

A **histogram** is a graphical representation of a frequency distribution. Bars are drawn over each class interval on a number line. The areas of the bars are proportional to the frequencies with which data fall into the class intervals.

The shape of a unimodal distribution of a quantitative variable may be **symmetric** (right side close to a mirror image of left side) or skewed to the right or left. A distribution is **skewed to the right** if the right tail of the distribution is longer than the left and is **skewed to the left** if the left tail of the distribution is longer than the right.



(tail on left) (tail on right)

Video #3: Histograms

http://www.learner.org/courses/againstallodds/unitpages/unit03.html

Fill in the Video Guide during the video.

Video Guide Answers:

1. The video opens by describing a study of lightning strikes in Colorado. What variable does the first histogram display?

Time of first lightning flash.

2. In this lightning histogram, what does the horizontal scale represent? What does the vertical scale represent?

2. Horizontal scale: Time of day in hours.
Vertical Scale: Percent of days with first lightning flash within that hour.

3. Was the overall shape of this histogram symmetric, skewed, or neither?

3. Roughly symmetric.

4. Why were a few values in the second lightning histogram called outliers?

4. These were values that were separated from the overall pattern by a gap in the data.

5. When you choose the classes for a histogram, what property must the classes have if the histogram is to be correct?

5. The classes need to have equal width.

6. What happens to a histogram if you use too many classes? What happens if you use too few?

6. Using too many classes can make it difficult to summarize patterns connected with specific values on the horizontal axis. (In other words, you can't see the forest for the trees.) Too few classes can mask important patterns.

Video Exercise 1

State	Total	65 and older	Percent 65 and older	State	Total	65 and older	Percent 65 and older
Alabama	4,780	658	13.80%	Montana	989	147	14.90%
Alaska	710	55	7.70%	Nebraska	1,826	247	13.50%
Arizona	6,392	882	13.80%	Nevada	2,701	324	12.00%
Arkansas	2,916	420	14.40%	New Hampshire	1,316	178	13.50%
California	37,254	4,247	11.40%	New Jersey	8,792	1,186	13.50%
Colorado	5,029	550	10.90%	New Mexico	2,059	272	13.20%
Connecticut	3,574	507	14.20%	New York	19,378	2,618	13.50%
Delaware	898	129	14.40%	North Carolina	9,535	1,234	12.90%
District of Columbia	602	69	11.50%	North Dakota	673	97	14.40%
Florida	18,801	3,260	17.30%	Ohio	11,537	1,622	14.10%
Georgia	9,688	1,032	10.70%	Oklahoma	3,751	507	13.50%
Hawaii	1,360	195	14.30%	Oregon	3,831	534	13.90%
Idaho	1,568	195	12.40%	Pennsylvania	12,702	1,959	15.40%
Illinois	12,831	1,609	12.50%	Rhode Island	1,053	152	14.40%
Indiana	6,484	841	13.00%	South Carolina	4,625	362	7.80%
Iowa	3,046	453	14.90%	South Dakota	814	117	14.40%
Kansas	2,853	376	13.20%	Tennessee	6,346	853	13.40%
Kentucky	4,339	578	13.30%	Texas	25,146	2,602	10.30%
Louisiana	4,533	558	12.30%	Utah	2,764	249	9.00%
Maine	1,328	211	15.90%	Vermont	626	91	14.50%
Maryland	5,774	708	12.30%	Virginia	8,001	977	12.20%
Massachusetts	6,548	903	13.80%	Washington	6,725	828	12.30%
Michigan	9,884	1,362	13.80%	West Virginia	1,853	297	16.00%
Minnesota	5,304	683	12.90%	Wisconsin	5,687	777	13.70%
Mississippi	2,967	380	12.80%	Wyoming	564	70	12.40%
Missouri	5,989	838	14.00%				

Table 3.3. Count (in Thousands) of people over 65 by State and the District of Columbia in 2010.

1. How many people in your state are at least 65 years old? The answer varies from state to state. Table 3.3 gives the data for all 50 states and the District of Columbia for the year 2010.

a. Make a histogram for these data. Use class intervals of width 500,000.

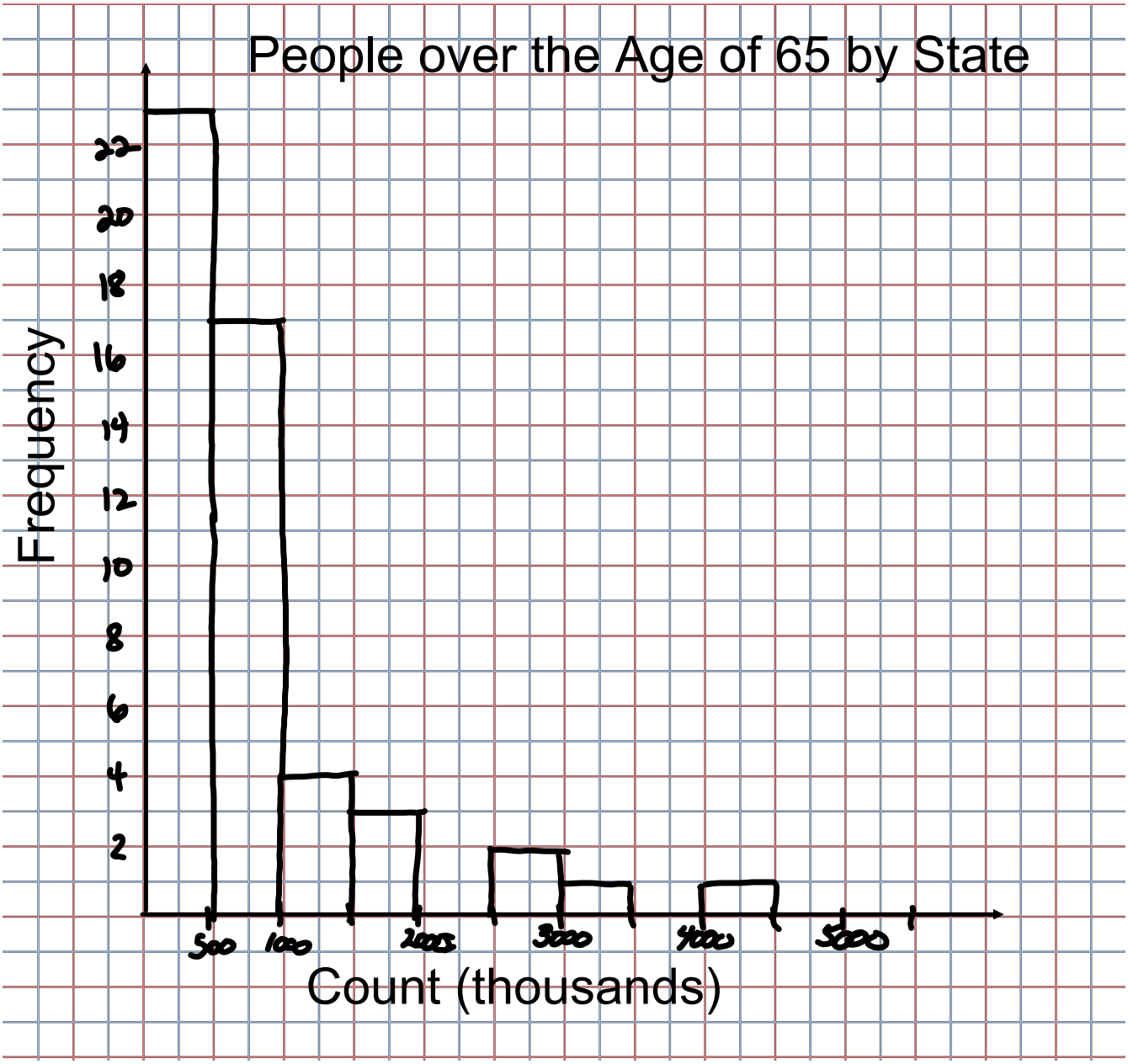
Excerise 1a. (Data from Table 3.3)

x axis

y axis

Count (in thousands)	Tally	Frequency
0-499	𝍸 𝍸 𝍸 𝍸 III	23
500-999	𝍸 𝍸 𝍸 II	17
1000-1499	IIII	4
1500-1999	III	3
2000-2499		0
2500-2999	II	2
3000-3499	I	1
3500-3999		0
4000-4499	I	1

Put graph on graph paper.


People over the Age of 65 by State
Frequency
22
20
18
16
14
12
10
8
6
4
2
500
1000
2000
3000
4000
5000
Count (thousands)