

T102 SECTION 8.1 STATISTICAL GRAPHS

For many years, the word *statistics* referred to numerical information about state or political territories. The word itself comes from the Latin *statisticus*, meaning “of the state”. We now live in an information age and the study of statistics is more important than ever before. In today’s world, much of statistics involves making sense of data.

Visual illustrations are an important way to depict information from statistics. These visual illustrations are simply pictures that display data – which may then tell us a story about the data.

In this section, we will be exploring the following visual depictions of information:

Pictographs	Line Plots	Stem and Leaf Plots
Frequency Tables	Grouped Frequency Tables	Histograms and Bar Graphs
Line Graphs	Scatterplots	Circle Graphs or Pie Charts

I. Pictograph - Page 505

A symbol or an icon is used to represent a certain quantity of items. A *legend* tells what the symbol represents.

II. Dot Plot (sometimes called line plot)

A line plot is similar to a pictograph, but we don’t lose the exact numerical data. This is a quick, simple way of organizing one set of numerical data, but typically with fewer than 50 values.

President	Age at Death	President	Age at Death	President	Age at Death
Washington	67	Fillmore	74	Roosevelt	60
Adams	90	Pierce	64	Taft	72
Jefferson	83	Buchanan	77	Wilson	67
Madison	85	Lincoln	56	Harding	57
Monroe	73	Johnson	66	Coolidge	60
Adams	80	Grant	63	Hoover	90
Jackson	78	Hayes	70	Roosevelt	63
Van Buren	79	Garfield	49	Truman	88
Harrison	68	Arthur	57	Eisenhower	78
Tyler	71	Cleveland	71	Kennedy	46
Polk	53	Harrison	67	Johnson	64
Taylor	65	McKinley	58	Nixon	81

A line plot for this data consists of a horizontal number line on which each score is denoted by an x above the corresponding value. Let’s construct a line plot for the first column of presidents:

Features that become apparent are:

OUTLIERS - data points whose value is significantly larger or smaller

GAPS - large space between points

CLUSTERS – an isolated group of points

III. Stem and Leaf Plot

****Don't forget the legends****

Similar to a line plot, but usually vertical and numbers are used instead of x's. To construct a stem and leaf plot for the president data, let's assign the ten's digit as the "stems" and the one's digits as the "leaves". (steps listed page 508)

Ages of Presidents at Death

4 |
5 |
6 |
7 |
8 |
9 |

Now let's make an **Ordered Stem and Leaf**
(Just arrange the leaves least to greatest)

A Back-to-Back Stem and Leaf
(Divide the group into early and late presidents)

Advantages of a Stem and Leaf Plot:

no original data lost
easy to create by hand
"similar to histogram" if turn sideways

IV. Frequency Table

A frequency table shows the number of times a certain piece of data occurs. Let's make one which describes the number of siblings we have:

Number of Siblings	Tally	Frequency
0		
1		
2		
3		
4		
5		
6		

V. Grouped Frequency Tables

Whenever you have a large set of data, you can put your measurements into groups, called *classes* or *intervals*. The stem and leaf plots naturally group scores into *classes*. The following classes were formed from the president data: 40-49, 50-59, 60-69, 70-79, 80-89, and 90-99. What is the size of each class? _____ Use these classes and the presidential death data to fill in the following grouped frequency table.

Ages at Death	Tally	Frequency
40 - 49		
Total		

SETTING UP A GROUPED FREQUENCY TABLE

To set up a grouped frequency table for a given set of data, sometimes you will be given the exact number and size of intervals to use and other times you will need to decide this for yourself. The following are instructions on creating a grouped frequency table so that it will accommodate all of your measurements:

1. Decide on the INTERVAL OR CLASS SIZE
 - a. Subtract the smallest data point from the largest data point
 - b. Divide this number by the number of classes or intervals you plan to have in your table
 - c. Round this UP to a convenient value. This is your INTERVAL SIZE.
2. Start your first interval with either a given value, your smallest data point, or some other convenient value (that will, of course, include your smallest value).
3. Add the interval size to this first value. This will be the lower limit for the SECOND INTERVAL! Continue this process until you have the desired number of intervals.
4. Fill in the UPPER LIMITS for each interval.

EXAMPLE: Construct a grouped frequency table for these 20 test scores, using the lowest score to start the first class (interval) and making 5 classes or uniform width.

31	30	23	27	19	26	28	38	17	29
26	34	21	23	23	22	12	26	39	25

VI. Histogram

We can look at a graphical representation of this data as well using a histogram. If you turn a stem and leaf plot sideways, you can see how a histogram is formed. The advantage that the stem and leaf has over the histogram is that you can see each actual data point whereas on the histogram you can only see the number of occurrences of each outcome. Using the same intervals as above, let's create a histogram. (We use a squiggle to indicate that part of the scale has been omitted; therefore the scale is not accurate at this area.)

Bar Graph

A Bar Graph is just like a histogram, but typically has spaces between the bars. Let's make a Bar Graph with our sibling data.



Double Bar Graph (Page 512)

This can be used to make comparisons between two sets of data. Use our back-to-back stem and leaf president data to construct a Double Bar Graph comparing the death ages between the early and late presidents.

VII. Line Graph

We use line graphs to show the **trend of a variable over time**. In fact, time is usually marked on the horizontal axis. Let's make a line graph that shows the trend of gasoline prices over the last 8 weeks:

Week 1: \$1.99	Week 2: \$1.69	Week 3: \$1.39	Week 4: \$1.59
Week 5: \$1.19	Week 6: \$1.09	Week 7: \$1.29	Week 8: \$1.79

VIII. Scatterplots

Given two sets of data we can use a scatterplot to determine a relationship between the data sets. This is generally done using ordered pairs and after graphing we can observe the data and see if a correlation exists. The points might fall along a "trend line". We can then use this to make predictions, but we cannot use it to deduce any cause and effect relationship.

IX. Circle Graphs or Pie Charts

This is a circular region partitioned into disjoint sections, with each section representing a part or percentage of the whole.

The disadvantage is that circle graphs are more challenging to construct.

For instance, we are given the following information:

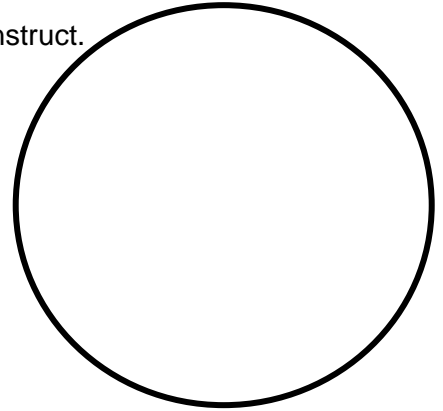
Of all hardrock miners interviewed

74% of them were white males

19% were white females

6% were nonwhite males

1% were nonwhite females



Descirption	PerCent	Degrees of circle
White males	74%	
White females	19%	
Nonwhite males	6%	
Nonwhite females	1%	

DATA FOR USE IN CLASS:

President	Age at Death	President	Age at Death	President	Age at Death
Washington	67	Fillmore	74	Roosevelt	60
Adams	90	Pierce	64	Taft	72
Jefferson	83	Buchanan	77	Wilson	67
Madison	85	Lincoln	56	Harding	57
Monroe	73	Johnson	66	Coolidge	60
Adams	80	Grant	63	Hoover	90
Jackson	78	Hayes	70	Roosevelt	63
Van Buren	79	Garfield	49	Truman	88
Harrison	68	Arthur	57	Eisenhower	78
Tyler	71	Cleveland	71	Kennedy	46
Polk	53	Harrison	67	Johnson	64
Taylor	65	McKinley	58	Nixon	81