Chapter 3 – Graphical and Tabular Displays of Data
Section 3 – Dotplots, Stemplots, and Time-Series Plots

Materials Needed:

Objectives
1 Identify discrete variables and continuous variables.
2. Construct and interpret dotplots.
3. Identify outliers of a distribution.
4. Find percentiles of a distribution.
5. Estimate the center of a distribution.
6. Construct and interpret stemplots.
7. Construct and interpret time-series plots.

Vocabulary
1. discrete variable
2. continuous variable
3. dotplot
4. frequency of an observation
5. frequency distribution of a numerical variable
6. outlier
7. *k*th percentile
8. stemplot/stem-and-leaf plot
9. split stem
10. time-series plot

Lesson/Activity
OBJECTIVE 1 − Identify discrete variables and continuous variables.

Definition: Discrete variable
A **discrete variable** is a variable that has gaps between successive, possible values.

Definition: Continuous variable
A **continuous variable** is a variable that can take on any value between two possible values.

Identify whether the variable is discrete or continuous.
1. the number of students in a prestatistics course

2. the volume (in gallons) of water in a lake

3. a person's height (in inches)

4. the price (in dollars) of a candy bar

When identifying a variable as discrete or continuous, we consider the possible values of the variable before rounding.

OBJECTIVE 2 − Construct and interpret dotplots.

To construct a **dotplot**, for each observation, we plot a dot above the number line, stacking dots as necessary.

5. The endorsements of the 15 athletes with the highest 2015 endorsements are shown in the following table.

Athlete	Sport	Endorsement (millions of dollars)
Tiger Woods	Golf	50
Cristiano Ronaldo	Soccer	27
Usain Bolt	Track and Field	21
Lebron James	Basketball	44
Kobe Bryant	Basketball	26
Kevin Durant	Basketball	35
Rory McIlroy	Golf	32
Rafael Nadal	Tennis	28
Mahendra Singh Dhoni	Cricket	27
Roger Federer	Tennis	58
Maria Sharapova	Tennis	23
Novak Djokovic	Tennis	31
Lionel Messi	Soccer	22
Phil Mickelson	Golf	44
Neymar da Silva Santos Junior	Soccer	17

Source: Opendorse

a. Construct a dotplot of the endorsements.

b. What observation(s) occurred the most?

c. How many observations are at least $40 million?

d. What proportion of the observations are at most $30 million?

We tend to use dotplots to describe data values of discrete variables, but they can be used to describe data values of continuous variables, too.

Definition: Frequency of an observation

The **frequency of an observation** of a numerical variable is the number of times the observation occurs in the group of data.

Definition: Frequency distribution of a numerical variable

The frequency distribution of a numerical variable is the observations together with their frequencies.

OBJECTIVE 3 – Identify outliers of a distribution.

An outlier is an observation that is quite a bit smaller or larger than the other observations.

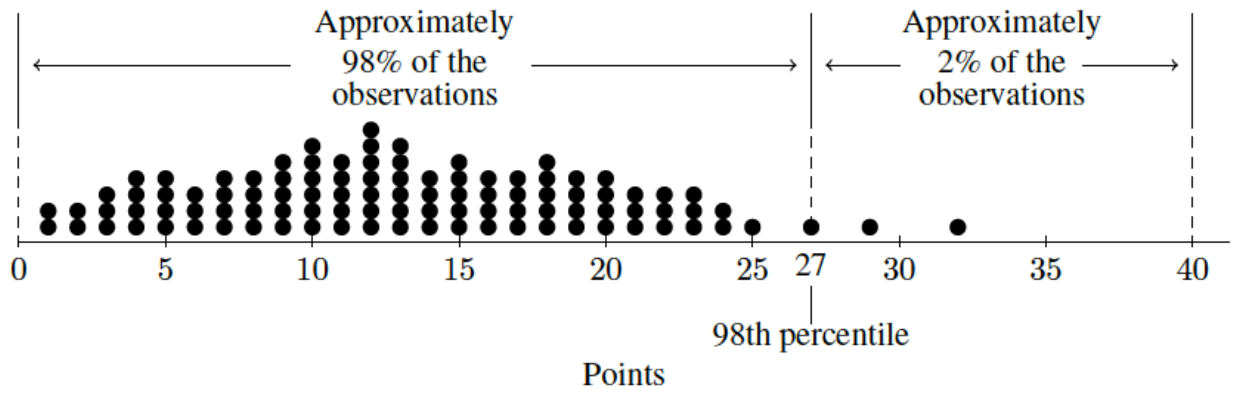
6. The maximum recorded lifetimes of animals that have the 10 largest maximum recorded lifetimes are shown in the following table.

Animal	Maximum Recorded Lifetime (years)
Ocean Quahog	400
Bowhead Whale	211
Rougheye Rockfish	205
Red Sea Urchin	200
Galapagos Tortoise	177
Shortracker Rockfish	157
Lake Sturgeon	152
Aldabra Giant Tortoise	152
Orange Roughy	149
Warty Orea	140

Source: Discovery News

a. Construct a dotplot.
b. Identify any outliers.

OBJECTIVE 4 – Find percentiles of a distribution.
For the data described by the following dotplot, 27 points is at the 98th percentile.



Definition: Percentile

The ***k*th percentile** of some data is a value (not necessarily a data value) that is greater than or equal to approximately k% of the observations and is less than approximately (100 – k)% of the observations.

7. For the endorsement data in Problem 5, find the percentile of Maria Sharapova.
8. For the endorsement data, find the 90th percentile.

OBJECTIVE 5 – Estimate the center of a distribution.
For this chapter, we will always use the 50th percentile to measure the center of a distribution.

9. For the endorsement data in Problem 5, use the 50th percentile to measure the center.

OBJECTIVE 6 – Construct and interpret stemplots.
A **stemplot** (or **stem-and-leaf plot**) breaks up each data value into two parts: the **leaf**, which is the rightmost digit, and the **stem**, which is the other digits.

10. Construct a stemplot for the endorsement distribution.

A **split stem** is a stemplot which lists the leaves from 0 to 4 in one row and lists the leaves from 5 to 9 in the next.

11. Construct a split stem by hand for the endorsement distribution.
12. Use StatCrunch to construct a split stem for the endorsement distribution.

Stemplots work best with a small number of observations whose variable is discrete or continuous with rounded values.

OBJECTIVE 7 – Construct and interpret time-series plots.
To construct a **time-series plot**, we plot points in a coordinate system where the horizontal axis represents time and the vertical axis represents some other quantity, and we draw line segments to connect each pair of successive
dots.

13. The number of new apps submitted to Apple's App Store per month are shown in the following table for various years.

Year	Number of Apps (thousands per month)
2009	5
2010	13
2011	13
2012	20
2013	25
2014	34
2015	40

Source: pocketgamer.biz

a. Construct a time-series plot of the data.
b. Describe how the number of new apps submitted to Apple's App store per month has changed over the years.
c. Find the change in the number of new apps submitted to Apple's App store per month from 2009 to 2015. If the number of new apps submitted to Apple's App store per month changes by that same amount from 2015 to 2021, what will it be in 2021? Do you have much faith that this will turn out to be true? Explain.

If a variable has increased (or decreased) throughout a period, we cannot assume it will continue to increase (or decrease) before or after that period.

Homework/Assessment
1, 3, 5, 7, 13, 15, 17, 21, 23, 25, 29, 31, 35, 37, 39