Pg. 10

What's the difference between categorical and quantitative variables?

Do we ever use numbers to describe the values of a categorical variable?

When describing the distribution of a quantitative variable, what characteristics should be addressed?

Alternate Example: US Census Data
Here is information about 10 randomly selected US residents from the 2000 census.

State	Number of Family Members	Age	Gender	Marital Status	Total Income	Travel time to work
Kentucky	2	61	Female	Married	21000	20
Florida	6	27	Female	Married	21300	20
Wisconsin	2	27	Male	Married	30000	5
California	4	33	Female	Married	26000	10
Michigan	3	49	Female	Married	15100	25
Virginia	3	26	Female	Married	25000	15
Pennsylvania	4	44	Male	Married	43000	10
Virginia	4	22	Male	Never married/ single	3000	0
California	1	30	Male	Never married/ single	40000	15
New York	4	34	Female	Separated	30000	40

(a) Who are the individuals in this data set?

(b) What variables are measured? Identify each as categorical or quantitative. In what units were the quantitative variables measured?

(c) Describe the individual in the first row.

Pg. 21-29

The "w's"

Counts

Area principle

Bar chart   Pie chart

When is it inappropriate to use a pie chart?

What are some common ways to make a misleading graph?

What is wrong with the following graph?

Read pg. 24 - 29

What is a two-way table? What is a marginal distribution?

What is a conditional distribution? How do we know which variable to condition on?

Pg. 30 - 35

What is a segmented bar graph? Why are they good to use?

What does it mean for two variables to have an association? How can you tell by looking at a graph?

Alternate Example: Super Powers
A sample of 200 children from the United Kingdom ages 9–
17 was selected from the CensusAtSchool website. The
gender of each student was recorded along with which
super power they would most like to have: invisibility,
super strength, telepathy (ability to read minds), ability to
fly, or ability to freeze time.

	Female	Male	Total
Invisibility	17	13	30
Super Strength	3	17	20
Telepathy	39	5	44
Fly	36	18	54
Freeze Time	20	32	52
Total	115	85	200

(a) Explain what it would mean if there was no association
between gender and superpower preference.
(b) Based on this data, can we conclude there is an association between gender and super power
preference? Justify.

Pg. 44

What is a distribution?

Pg. 45

What is the difference between a frequency table and a relative frequency table? When is it better to use relative frequency tables?

Pg. 44-45

Histograms   (half-day)

The following table presents the average points scored per game (PPG) for the 30 NBA teams in the 2009–2010 regular season. Make a dotplot to display the distribution of points per game. Then, use your dotplot to make a histogram of the distribution.

Team	PPG	Team	PPG	Team	PPG
Atlanta Hawks	101.7	Indiana Pacers	100.8	Oklahoma City Thunder	101.5
Boston Celtics	99.2	Los Angeles Clippers	95.7	Orlando Magic	102.8
Charlotte Bobcats	95.3	Los Angeles Lakers	101.7	Philadelphia 76ers	97.7
Chicago Bulls	97.5	Memphis Grizzlies	102.5	Phoenix Suns	110.2
Cleveland Cavaliers	102.1	Miami Heat	96.5	Portland Trail Blazers	98.1
Dallas Mavericks	102	Milwaukee Bucks	97.7	Sacramento Kings	100
Denver Nuggets	106.5	Minnesota Timberwolves	98.2	San Antonio Spurs	101.4
Detroit Pistons	94	New Jersey Nets	92.4	Toronto Raptors	104.1
Golden State Warriors	108.8	New Orleans Hornets	100.2	Utah Jazz	104.2
Houston Rockets	102.4	New York Knicks	102.1	Washington Wizards	96.2

How do you make a histogram?

Why would we prefer a *relative* frequency histogram to a frequency histogram?

Pg. 47-48

What is the most important thing to remember when making a stemplot?

Alternate Example: Which gender is taller, males or females? A sample of 14-year-olds from the United Kingdom was randomly selected using the CensusAtSchool website. Here are the heights of the students (in cm). Make a back-to-back stemplot and compare the distributions.

   Male: 154, 157, 187, 163, 167, 159, 169, 162, 176, 177, 151, 175, 174, 165, 165, 183, 180

   Female: 160, 169, 152, 167, 164, 163, 160, 163, 169, 157, 158, 153, 161, 165, 165, 159, 168, 153, 166, 158, 158, 166

Pg. 49

What is the most important thing to remember when you are asked to compare two distributions?

Alternate Example: Energy Cost: Top vs. Bottom Freezers

How do the annual energy costs (in dollars) compare for refrigerators with top freezers and refrigerators with bottom freezers? The data below is from the May 2010 issue of *Consumer Reports*.
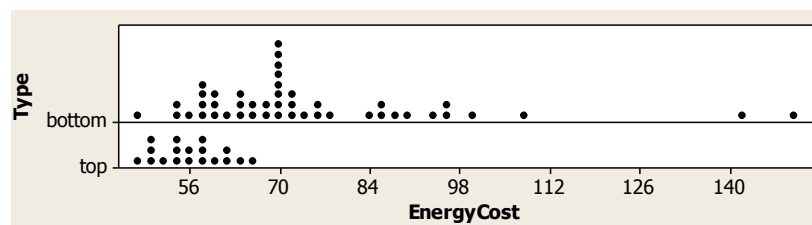
Displaying Quantitative Data with Graphs

Pg. 49 – 50

When describing the distribution of a quantitative variable, what characteristics should be addressed?

Briefly describe/illustrate the following distribution shapes:
Symmetric   Skewed right   Skewed left

Unimodal   Bimodal   Uniform

Alternate Example: Smart Phone Battery Life
Here is the estimated battery life for each of 9 different smart phones (in minutes). Make a dotplot of the data and describe what you see.

Smart Phone	Battery Life (minutes)
Apple iPhone	300
Motorola Droid	385
Palm Pre	300
Blackberry Bold	360
Blackberry Storm	330
Motorola Cliq	360
Samsung Moment	330
Blackberry Tour	300
HTC Droid	460

Describing Quantitative Data with Numbers

Pg. 52 - 53

Is the median a resistant measure of center? Explain.

How does the shape of a distribution affect the relationship between the mean and the median?

Pg. 54

What is the range? Is it a resistant measure of spread?   Explain.

What are quartiles? How do you find them?

What is the interquartile range (*IQR*)? Is the *IQR* a resistant measure of spread?

Pg. 59

What is the difference between $\bar{x}$ and μ ?

.

What is a resistant measure? Is the mean a resistant measure of center?

How can you estimate the mean of a histogram or dotplot?