The Data

Gender	Texting style	Left thumb length	Right thumb length	A total score	B total score	Total both scores	A minus B	Avg Thumbs	Diff Thumbs
Male	Char	6.5	6.5	35	25	60	10	6.5	0
Female	Char	5	5	61	57	118	4	5	0
Female	Word	6	6	24	20	44	4	6	0
Male	Word	7	7	43	60	103	-17	7	0
Female	Word	6	6	14	19	33	-5	6	0
Male	Word	7	6	15	18	33	-3	6.5	-1
Female	Word	6	6	13	14	27	-1	6	0
Female	Word	6	6	22	10	22	12	6	0
Male	Word	6.5	6	13	15	28	-2	6.25	-0.5
Female	Word	5.5	5.5	16	16	32	0	5.5	0
Male	Char	6	5	85	78	163	7	5.5	-1
Male	Char	6	6	126	120	246	6	6	0
Male	Word	7.5	6.5	67	69	136	-2	7	-1
Female	Char	5.5	5.5	11	7	18	4	5.5	0
Female	Word	5.5	5.7	14	17	31	-3	5.6	0.2
Female	Word	6	6	17	14	31	3	6	0
Female	Word	5	5	20	15	35	5	5	0
Male	Word	6.5	6.5	15	13	28	2	6.5	0
Male	Word	7	7	30	31	61	-1	7	0
Male	Word	6	6.1	120	117	237	3	6.05	0.1
Male	Word	6	6	74	25	99	49	6	0
Male	Word	6.3	6	23	21	44	2	6.15	-0.3
Male	Char	6	5.9	45	50	95	-5	5.95	-0.1
Male	Word	6	6.1	86	100	186	-14	6.05	0.1
Male	Char	6	6	25	23	48	2	6	0
Male	Char	6.3	6.3	81	57	138	24	6.3	0
Female	Char	5.5	5.5	88	66	154	22	5.5	0
Female	Word	7	6.8	10	9	19	1	6.9	-0.2
Male	Char	6.5	7	21	18	39	3	6.75	0.5
Female	Char	5.4	5.2	72	48	121	24	5.3	-0.2
Female	Char	8	8	36	23	59	13	8	0
Male	Char	7	6.5	46	45	91	1	6.75	-0.5
Female	Char	6	6.8	48	39	87	9	6.4	0.8
Female	Char	7.1	7.1	84	57	141	27	7.1	0
Female	Word	5.9	5.5	25	23	48	2	5.7	-0.4
Female	Char	7.6	7.2	32	45	77	-13	7.4	-0.4
Male	Word	6.9	7	23	28	51	-5	6.95	0.1
Female	Char	7.7	7.5	18	15	33	3	7.6	-0.2
Male	Char	8.6	(cast)	22	20	42	2	N/A	N/A
Male	Char	7.3	7.1	54	50	104	4	7.2	-0.2

Text Messaging: Communication in the Twenty-First Century

In the natural world there are many different forms of communication between and among species of flora and fauna. Explorer ants leave scent trails for worker ants; bees distribute pollen from flower to flower, facilitating fertilization. From bright colors to camouflage, and from loudly voiced noises to quiet movements, every species has its own unique form of communication.

Beginning in the late nineteenth century and continuing through much of the twentieth century, Morse code was one of the ways used by the human species to send important messages. When transmitting telegraphic information, standardized sequences of short and long elements were used to represent the letters, numerals, punctuation, and special characters of a message. The short and long elements are formed by sounds, marks, or pulses in "on-off" keying and are commonly known as "dots" and "dashes" or "dits" and "dahs."

In recent years a new form of communication has arisen that connects people—especially the teenage subspecies—in a powerful way: text messaging! Text messaging, also called SMS (Short Message Service) allows short text messages to be sent and received on a mobile phone. Messages can be sent from one phone to another by addressing the message to the recipient's phone number. Messages can also be sent to a phone via a special email address, through the carrier's website, or with special messaging software and a modem. Most phones and carriers also allow messages to be sent from a phone directly to an email address.

Text messaging as a form of communication is still developing, as is the language that defines and surrounds it. Text messengers frequently use abbreviations both for speed and cost-efficiency. Speaking of speed, a 13-year-old girl recently won a $25,000 prize at a text-messaging speed competition in New York. Another competition took place on Jay Leno's *The Tonight Show*, pitting a young, male champion cell-phone texter against an elderly ham-radio buff with a telegraph and the 140-year-old technology of Morse code. (We won't tell you who won—to find the video, do a Google search for "text messaging Leno.") These events have brought national attention to this new medium of communication and a whole slew of interesting questions ripe for possible data analysis.

Some questions that might be asked are:

- Do males and females differ in their text-messaging style or speed?
- What variables might be used to predict the speed at which an individual can text? Experience? Phone provider? Thumb length? Text-messaging style?
- What do the distributions of variables associated with the text-messaging experience show us? For example, how about the number of texts that students send in one day? Or the accuracy with which they can text?
- Do adults stand a chance when compared to a high school texting expert?

To address and potentially answer these and other questions, data are needed. We have constructed a survey that can be used to help you collect your own data, and directions for administering the survey are provided so that the data can be collected consistently. We have also collected data from individuals at some high schools in southern California and Indiana (admittedly a convenience sample). If your teacher chooses to use these data, you will be shown how to access it. Irrespective of whether you gather your own data or use ours, we trust that your analyses will be enlightening. We will help you through the analysis of these data by asking some well-chosen questions—your part of this data analysis activity is providing some well-chosen responses!

Part 1

"Is it a boy or a girl?" This may be the first question people ever asked about you (perhaps even before your birth!). Much research has been conducted in a variety of different areas to try to answer questions about gender differences. Let's join the parade and bring some questions of gender into text messaging. Does gender influence text-messaging habits? Do females text faster than males? Do males have longer thumbs than females? Let's see if the data can help answer some of these questions related to gender and other categorical variables.

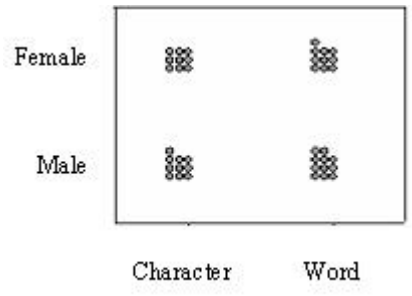
Question 1: Some people may believe that girls have better communication skills than boys. But are females faster "texters" than boys? Consider the "Total both scores" data. (The TBS variable is found by summing the scores from both sentences.)

1a. Describe two appropriate graphic methods that could be used to show any gender differences in total score. Describe the advantages and disadvantages of each graph.

1b. Create one of the two graphs you described in part 1a and sketch it below. Write a few sentences, discussing your conclusions based on the graphs. Supply statistical justification for your conclusions.

Question 2: Does gender affect the type of texting used (character vs. word)?

2a. A graph called a breakdown plot is shown below. It shows the number of students in each of four categories. What general conclusion can you make regarding gender and texting style?



Breakdown Plot Texting Style

2b. Consider the table below of counts of students who used character and word styles. In light of these overall totals and the individual cell counts, does it appear that texting style is **independent** of gender? Explain.

	Texting Style		
	Character	Word	
Female	9	10	**19**
Male	10	11	**21**
	19	**21**	**40**

Part 2

Time is a valuable commodity and saving time is one of the important aspects of text messaging. In part 2, we will examine some variables that seem to be associated with the texting speed of the subjects.

In general terms, we might think of a message that takes longer to text or has more mistakes in the result as a more "difficult" passage to text. It is this concept of difficulty that we will explore in part 2. The sentences from the survey are presented below.

Question 1: Is sentence A more difficult to text than sentence B?

A: Statistics students are above average.
B: Meet me at my car after school today.

1a. What are some variables that might indicate any difference in "difficulty"—variables that could be investigated using the data?

1b. First, let's examine the time it took to type each message, ignoring the mistakes made. Make a graphical display that compares the times for the two sentences. Write a few sentences about what you observe, consulting your graphical display.

1c. Now make a graph that compares the number of mistakes made on each of the sentences. Describe in a few sentences any differences in the distributions, and (of course!) comment on any outliers.

1d. Our last comparison of the separate gender scores will be of the total scores for each of the two sentences. Again, in a few sentences describe what you see.

1e. Now calculate the difference of scores (A minus B) for each student. What is the mean of the differences?

Mean of differences (A minus B) = _____ seconds

1f. If you were assessing whether sentence A were more difficult than sentence B, how would you use the mean difference?

1g. About how large of a mean difference would convince you that sentence A *was* more difficult than sentence B? What reasoning led you to this number?

Question 2: We will now consider the idea of speed in texting. Is texting by word (using predictive-text software like T9) faster than texting by character? How might gender be associated with texting style and mean speed of texting? Do any other variables affect speed?

2a. In order to compare, for example, the total score of subjects using word style vs. the total score of subjects using character style, one might use different types of graphs. Choose one type of graph and compare the two distributions. Compute summary statistics for the word group and the character group.

2b. Find the mean total scores of each gender in each style:

Mean total score of female (word style) = _____ seconds

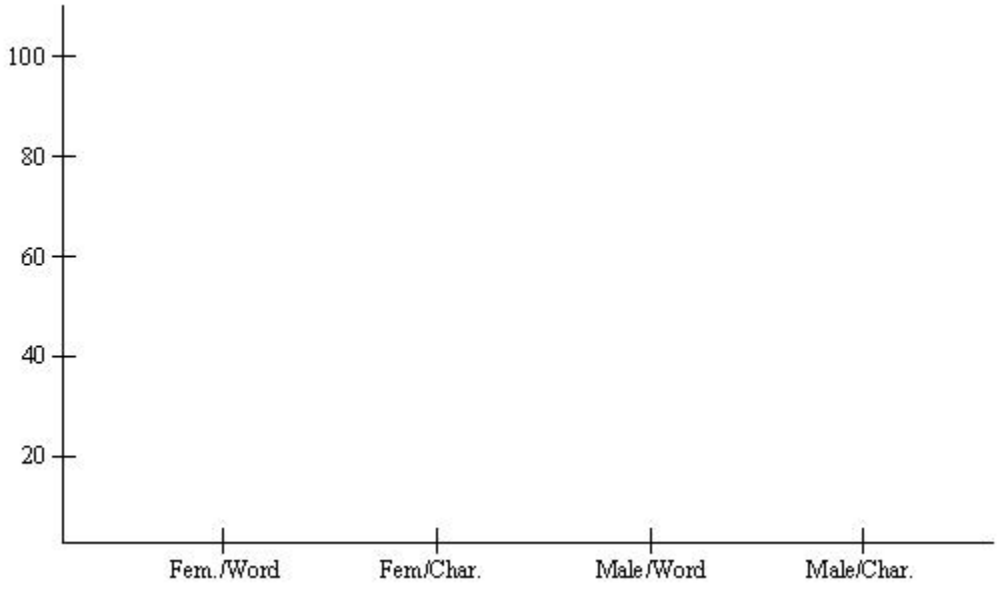
Mean total score of male (word style) = _____ seconds

Mean total score of female (character style) = _____ seconds

Mean total score of male (character style) = _____ seconds

2c. Does there appear to be a significant difference in mean total scores of the texting styles? Justify your answer.

2d. Now plot each mean on the graph below with an "X" indicating the mean for each group.



Mean Total Texting Time (in Seconds) by Gender and Style

2e. In a paragraph or two, describe and interpret any differences shown by the graph in part 2d.

Question 3: It might be that thumb length is different for boys and girls, and this difference in thumb length might explain why they might differ in texting variables. We now consider the question of thumb length: how is thumb length related to gender?

3a. Construct a graph to compare the thumb lengths of males and females. Do males have larger thumbs in general than females? Justify your answer by referring to specific aspects of your graph.

3b. Calculate all the **differences** in thumb lengths ("Right thumb length" minus "Left thumb length"). Graph the differences by gender, and describe the results.

3c. What is unusual about the box plot of differences in thumb length for females? (Hint: think about skewness.)

3d. Make a graphical display that compares left and right thumb length. Are our subjects' right and left thumbs always the same length? If there is a difference, what would you attribute this difference to?

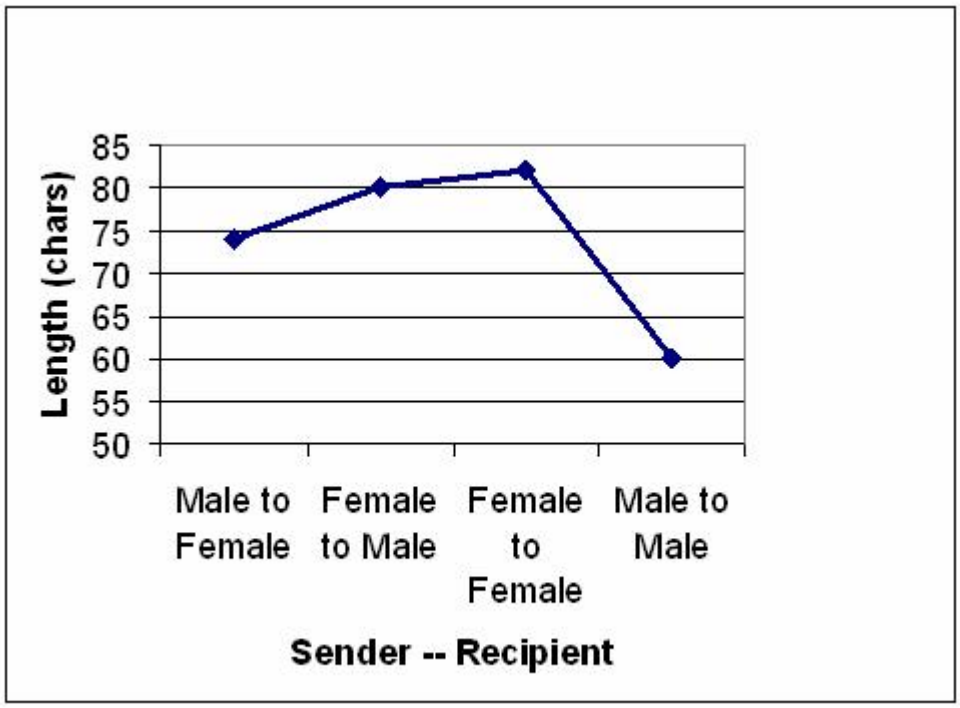
Question 4: (Extension) Do females send longer text messages than males?

Dr. Simeon Yates, a senior lecturer at Sheffield Hallam University in the United Kingdom, has studied texting communication and gender. Data from one of his presentations are given below:

Average Length (Characters) in Text Messages

Men texting to women: 74
Women texting to men: 80
Women texting to women: 82
Men texting to men: 60

A plot of these data is given below:



Mean Text Lengths by Gender Relation

4a. What conclusions do you draw from the graph? Justify your conclusions by referring to specific aspects of the graph.

Part 3

In part 3 we will consider the individual responses to the two sentences, and how the total scores for the two sentences are related.

1a. We will start by comparing sentence A total score to sentence B total score for all of the subjects. Use a calculator or computer to create a scatter plot of the data. (Put the sentence B score on the vertical axis.) Describe the features you see.

1b. Compute and **interpret** r and r^2 for this data.

1c. Compute the least-squares regression line for this data. How would you interpret the intercept and the slope in this context? Upon reflection and reconsideration, do your interpretations "make sense" in the real world?

1d. Make a residual plot of your data. Identify and discuss any outliers and/or influential points.

1e. Predict the scores for text message B for someone who had a score of 22 seconds + errors and someone who had a score of 5 seconds + errors. Which prediction (if either) do you feel is more reliable, and why?

2. Does thumb length predict texting score? Repeat questions 1a through 1e with this data to answer the question.

3. Select another pair of variables from the data that are appropriate for bivariate analysis. Explore the relationship between the variables you select using questions 1a through 1e above.