1.   Here are two statistics on National Football League players in 1997

$500,000 per year
$1,750,000 per year

Which one of these numbers is the mean salary and which one is the median salary of National Football League Players?

The mean is 1,750,000

The median is 500,000

Explain your choice in the space below.

Salaries can range from zero to positive infinity but can't be negative. So chances are that the distribution has a long right handed tail which will pull the mean above the median.

2.   The IQ scores of adult humans (age 18 and over) is approximately normal with a mean of 100 and a standard deviation of 15.

(a)   How low is the lowest 5% of all IQ scores (that is, at or below what IQ score is the lowest 5%)
         How high is the highest 10% of IQ scores (that is, at or above what IQ Score is the highest 10%)?

Use a Z = -1.65 for lowest
Use a Z = +1.3 for the highest

Solve for X in each case using the formula:

   $Z = (X - 100)/15$

Should get 75.25 for lowest, 119.5 for highest

(b)   A simple random sample of 25 college students is drawn from the adult human population. The sample average is 108 and the sample standard deviation is 30. Please test the hypothesis that college students have higher IQ scores than the average human. State a null and an alternative hypothesis, perform a test, state a p-value and explain your result (do you reject or not reject the null and why). Use a 5% level of significance as your decision rule.

The null is any difference in the IQ score is due to chance alone, the true mean is actually 100.

The alternative is that any difference in IQ is NOT due to random chance and the true mean is actually greater than 100.

$$Z = \frac{108 - 100}{15/\sqrt{25}} = 8/3 = +2.67 \text{ round to } 2.70$$

p-value is (100 - 99.31)/2 = .345% or .35% and this is much less than 5%

Reject the null at the 5% level, college students have higher Iqs on average

3.   Investors ask about the relationship between returns on investments (the money you make by investing your money) in the United States and on investments overseas. Below is a table of total returns on investments on U.S. and overseas stocks over a 10 year period.

	Year	Overseas % Return	U.S. % Return
	1987	24.6	5.1
	1988	28.5	16.8
	1989	10.6	31.5
	1990	-23	-3.1
	1991	12.8	30.4
	1992	-12.1	7.6
	1993	32.9	10.1
	1994	6.2	1.3
	1995	11.2	37.6
	1996	6.4	23
Average	1991.5000	9.8100	16.0300
Standard Deviation	2.7386	15.6493	12.6810

(a)   Find the correlation, r, of the U.S. and overseas returns then describe the relationship between U.S. and overseas returns in words, using r to make your description more precise.

I calculated .3239, weakly positive relationship suggesting that overseas returns and US returns tend to increase together, but the relationship is not strong.

(b)   Find the regression line of overseas returns on U.S. returns. Please interpret the values of the slope and of the intercept of this line.

Overseas returns = 3.403 + .3997(US returns)

(c)   (continued from above) In 1993, the return on U.S. stocks was 10.1%, what was the predicted return on overseas stocks. Is the predicted return the same as the actual return? If it is the same, please explain why this is so. If it is different, please explain why they are different.

Input 10.1 into the regression line above and get a new over seas return. It should not be the same. Review the section on residuals.

A study of the effects of running on mood involved 321 adult male runners who each ran an average of 20 miles per week. The runners were given a psychological test to measure their levels of happiness, sadness, anger, etc. A newspaper article stated, "The researchers found statistically significant mood differences between the runners and the adult male population as a whole." The headline on the article was "RESEARCH SHOWS THAT RUNNING CAN ALTER MOOD."

(a) Your best friend has not taken any statistics courses and knows nothing about statistics. Please tell me how you would explain to him or her what "statistically significant" means.

I think you can do this.

(b) Please also tell me how you would explain to your friend why the headline is misleading.

5. You got a job working for a marketing company and your supervisor is planning a sample survey of households in Los Angeles. Your supervisor instructs you to contact households by random-digit dialing phone numbers. Your supervisor knows from past experience that about 70% of the households you contact in this manner will respond.

(a) If you randomly dial 1500 telephone numbers, what are the mean and standard error of the number of households who respond?

Expect 1,050 calls because you get 70% of 1500. So the mean is 1,050 and the SE is:

$$SE = \sqrt{1500} * (1-0)\sqrt{.7 * .3} = 17.75$$

(b) Find the probability that you will get at least 1000 responses.
Z=(1000-1050)/17.75 = 2.82 round to 2.85

Probability is 99.56+(100-99.56/2) = 99.78%

6. A psychologist has a theory about what kinds of situations people find pleasant or unpleasant. To test his theory, he made up 100 questions of the following kind:

Suppose you like John, you like dogs and John likes dogs. How pleasant is this situation?

For each question, the subject is asked whether he or she likes some person (e.g. John) and then some object, and if so, whether the person likes the object as well. The psychologist presents the 100 questions to a volunteer subject, one question at a time. After each question, the subject rates the pleasantness of the situation on a scale of 1 to 5, with a 1=unpleasant, a 3 = neutral, and a 5=pleasant. The average of the ratings is 4.2 with a standard deviation of 0.3.

For the two questions below, first indicate whether the statement is true or false. Then explain your reasoning. If you need more information to decide whether it is true or false, say what you need and why you would need it to answer the question.

(a) An approximate 95% confidence interval for a subject's long-run average response to this kind of questioning is in the interval from 3.6 to 4.8.

TRUE. Show how to contruct the CI given the information above.

(b) Approximately 68% of the questions had responses in the range 4.2 ± 0.3
FALSE. Wrong reasoning. We cannot say this because we don't the composition of the box and we do not know if it is normally distributed. This is a mixed up statement, partly confidence interval, partly a statement of chance from Chapter 17 or 20.

7. You are planning to perform a significance test of

H_0: mean = 0

Versus

H_1: mean < 0

What values of Z would lead you to reject H_0 at the 1% level of significance? Then answer this question: True or False and explain why. A significance test that is significant at the 1% level of significance must always be significant at the 5% level of significance.

In this case, a Z less than -2.35 would be sufficient.
True.
1% will always be less than 5% check the table A 105.

8.   A study gives the following results for the law school admissions test (LSAT) and undergraduate GPA (grade point average):

Average GPA: 3.3;   Standard deviation   = 0.4
Average LSAT: 150;   Standard deviation   = 10
Correlation coefficient = 0.6

(a) Suppose one particular student has a GPA of 3.4, can you predict what this person got on the LSAT? Answer True or False and justify your answer.

TRUE. Find a and b from Chapter 10 and substitute, you should get something like 151.5

(b) Suppose a group of students all have GPA's of 3.4, can you predict what their average LSAT score will be? Answer True or False and justify your answer.

FALSE. Review Chapter 10.1 and Exercise Set A #3