AP Statistics – Summer Reading Assignment

Based upon your experiences with statistics in your previous math courses, you should be familiar with some of the basics. The following assignment reviews some very basic information about statistical variables and their nature, some information about the collection of data, linear regression, and some basic probability. If you don't remember this information from your previous classes or for some reason haven't seen it before, you can research the information online quite easily. One fairly useful resource is the AP Statistics tutorial at StatTrek: http://www.stattrek.com/AP-Statistics-1/AP-Statistics-Intro.aspx?Tutorial=ap# . There are several other reasonable sources available on the internet.

The assignment is due on the first Monday of the new school year, August 8th. Please contact Ms. Sikes at csikes@forsyth.k12.ga.us or Mrs. Graham at tgraham@forsyth.k12.ga.us if you have *any* questions regarding the assignment.

Note: Justify your answers. For example, when a question asks if a particular example of data is qualitative of quantitative, state your answer and then state *why* the data are quantitative or categorical. Show the instructor your reasoning process for arriving at your answer.

1. Identify each variable as quantitative or categorical:

   a.) Amount of time it takes to assemble a simple prize

   b.) Number of students in a first-grade classroom

   c.) Rating of a newly elected politician (excellent, good, fair, poor)

   d.) State in which a person lives

2. Identify the following quantitative variables as discrete or continuous:

   a.) Population in a particular area of the United States

   b.) Weight of newspapers recovered for recycling on a single day

   c.) Time to complete a sociology exam

   d.) Number of consumers in a poll of 1000 who consider nutritional labeling on food products important

3. A medical researcher wants to estimate the survival time of a patient after the onset of a particular type of cancer and after a particular regimen of radiotherapy.

   a.) What is the variable of interest to the medical researcher?

   b.) Is the variable in part A qualitative, quantitative discrete, or quantitative continuous?

4. Fifty people are grouped into four categories—A, B, C, and D—and the number of people who fall into each category is shown in the table:

Category	Frequency
A	11
B	14
C	20
D	5

a.) What is the variable being measured? Is it qualitative or quantitative?

5. The 1960s generation was never radical as it was portrayed. According to an opinion poll in *The American Enterprise*, when a group of 30-40-year-olds were asked to describe their political views in the 1960s and early 1970s, they gave these responses:

Conservative: 28%
Moderate: 35%
Liberal: 31%
Radical: 6%

Source: Karlyn Bowman, ed., "Opinion Pulse: '60s Kids: The Way They Were," *The American Enterprise,* May/June 1997; p. 91.

a.) Is the variable qualitative or quantitative?

6. This July, the U.S. House of Representatives voted to cut taxes for American citizens.

a). How would you classify the record of the vote in the 435-member House; is it a descriptive statistical study or an inferential statistical study?

b). A poll asks 435 citizens whether they want the tax cut bill passed. Is this study descriptive or inferential?

7. *National Geographic Magazine* (July, 1999) published an article called "The Shrinking World of Hornbills." (Hornbills are a genus of birds including 54 species.) Suppose you're a naturalist studying these birds. Please answer the following questions about your study of hornbills.

a.) If you were to study the total number of eggs produced in one month by female Red-Knobbed Hornbills, would your variable (number of eggs) be categorical, discrete quantitative, or continuous quantitative?

b.) If you wanted to know the weights of the eggs produced by Red-Knobbed Hornbills, would that variable (weight) be categorical, discrete quantitative, or continuous quantitative?

8. Here is a small part of a data set that describes Major League Baseball players as of opening day of the 2005 season:

Player	Team	Position	Age	Height	Weight	Salary
Ortiz, David	Red Sox	Outfielder	29	6-4	230	5,250,000
Nix, Laynce	Rangers	Outfielder	24	6-0	200	316,000
Perez, Antonio	Dodgers	Infielder	25	5-11	175	320,500
Piazza, Mike	Mets	Catcher	36	6-3	215	16,071,429
Rolen, Scott	Cardinals	Infielder	30	6-4	240	10,715,509

a.) What individuals does this data set describe?

b.) Identify the variables that were recorded. Label each as categorical or quantitative.

9. Students in a high school statistics class were asked to report how many siblings they have. Here are the data:

1	2	1	2	3	2	1	1	2	3
0	1	4	2	1	2	1	7	0	1
0	2	2	5	2	1	1	0	1	1
3	1	2	1	1	2	3	2	4	2

a.) Construct a dotplot to display these data.

b.) How many siblings does a "typical" student in the class have? Justify your answer.

c.) Are there any really unusual values in the data set? Explain.

10. A study of computer-assisted learning examined the learning of "Blissymbols" by children. Blissymbols are pictographs (think of Egyptian hieroglyphs) that are sometimes used to help learning-impaired children communicate. The researcher designed two computer lessons that taught the same content using the same examples. One lesson required the children to interact with the material, while in the other the children controlled only the pace of the lesson. Call these two styles "Active" and "Passive." Children were assigned at random to Active and Passive groups. After the lesson, the computer presented a quiz that asked the children to identify 56 Blissymbols. Here are the numbers of correct identifications by the 24 children in the Active group and the 24 children in the Passive group:
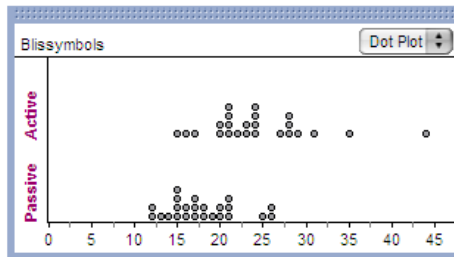
a.) Is the study described an observational study or an experiment? Explain.

b.) Describe any similarities and differences you see in the dotplots. Be specific.

c.) Do you think there is good evidence that active learning is superior to passive learning? Explain.

11. A student at a large high school conducts a study using a random sample of 20 male seniors. The student records the height (in inches) and the length of the right foot (in millimeters) of each student selected, and uses a calculator to fit a least squares regression line to the data. The calculator provided the following information:

predicted foot length = 148.17 + 1.8849*(height), $r^2 = .357$

a) What is the value of the correlation coefficient (r) for foot length and height?

b) A scatterplot of the results is shown below. Draw the least squares regression line on the scatterplot.
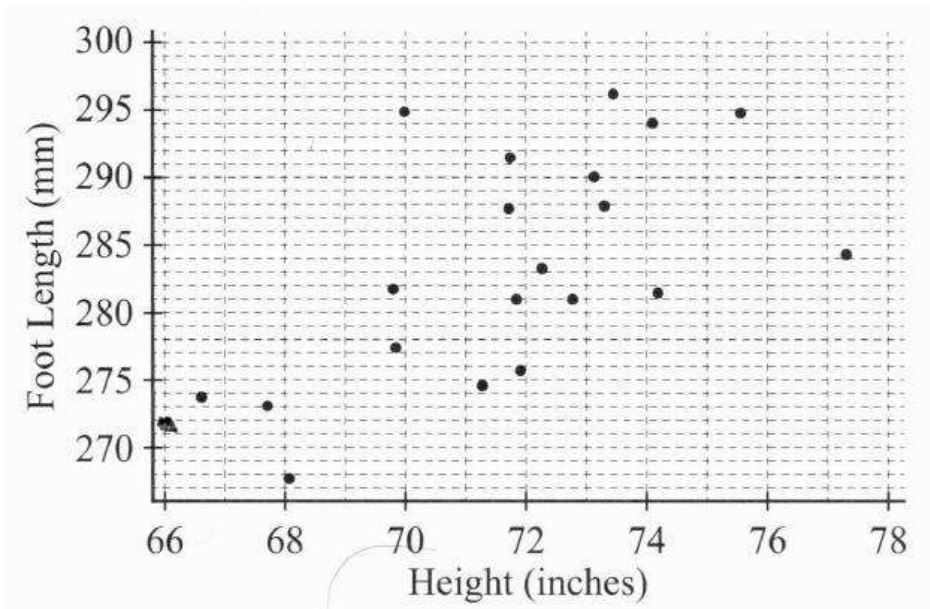
c) Interpret the value of the slope of the least squares regression line in this context.

d) Does the intercept of the least squares regression line have a meaningful interpretation in this context? If so, provide this interpretation. If not, explain why not.

e) What does the least squares regression line predict for the foot length of a student whose height is 73 inches?

12. A music enthusiast has a collection of recorded music consisting of 3390 albums. Each album has been classified as one of classical, jazz, or popular, and is on either CD, vinyl, or cassette. (There is only one recording of each album in the collection.) The numbers of albums falling into these categories are given in the table below.

	CD	Vinyl	Cassette	Total
Classical	846	690	81	1617
Jazz	693	562	116	1371
Popular	158	95	149	402
Total	1697	1347	346	3390

An album is chosen at random from this collection. Find the probability that it is:

a) Jazz

b) Jazz or on vinyl

c) Jazz and on vinyl

d) Jazz, given the condition that it is on vinyl

13. A student is about to take AP's in US History, English Language, and Statistics. She estimates that her probabilities of getting 5's in these subjects are 0.6, 0.7, and 0.8, respectively. She is also willing to assume that her results in the three subjects are independent. Assuming that the student's estimates are correct, find the probability that she gets:

a) 5's in all three subjects

b) No 5's

c) At least one 5

14. In a particular population of polar bears, the adult males have masses that are normally distributed with mean 515 kilograms and standard deviation 88 kilograms.

a) And adult male is chosen at random from this population. Calculate the probability that his mass is

   i.   Between 480 and 580 kilograms

   ii.   Less than 600 kilograms

   iii. More than 450 kilograms

b) What is the minimum mass required for a bear to be amongst the heaviest 20% of adult males in this population?