

In-class test _____/80 points

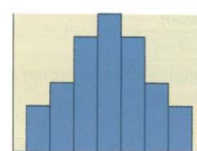
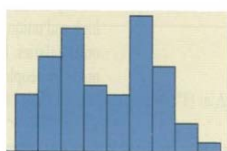
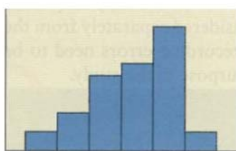
Take home test _____/20 points

1. Most breast cancer patients are over the age of 50 at diagnosis. A researcher at a particular New York cancer center believes that her patients are even older than the norm, typically older than 65 years at diagnosis. To investigate, she reviews the ages of a random sample of 100 of her patients diagnosed with breast cancer. Identify the following:

- (a) Population
- (b) Sample
- (c) Sample size
- (d) Variable of interest: _____
- (e) The variable of interest is (circle one) CATEGORICAL NUMERICAL

2. Choose from the list below to best describe the shape of each distribution:

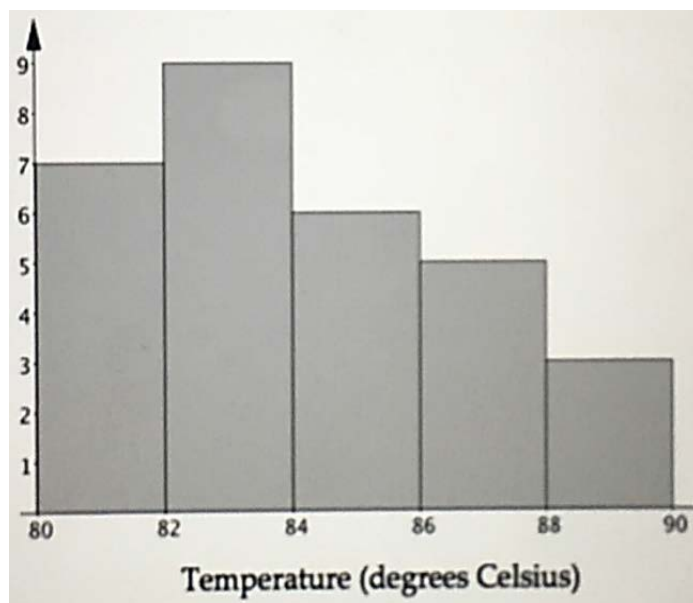
skewed left, skewed right, uniform, symmetric, unimodal (may apply to more than 1 graph), bimodal



3. (12 pts) The daily high temperatures (in degrees Fahrenheit) were recorded in SLO over a period of time in the summer. The histogram shows the distribution of temperatures over those days.

Use the histogram to answer the following questions.

- (a) How many days were in this study? _____
- (b) How many days had a high that was 86 degrees or more? _____
- (c) What is the relative frequency (express as a percent) of the days that had a high that was 86 degrees or more.
- (d) Which of the following could be the median and mean of the data? (circle the best answer...no work needed)



- (i) Median = 83.5, Mean = 83.5
- (ii) Median = 83.5, Mean = 81.5
- (iii) Median = 81.5, Mean = 83.5
- (iv) Median = 83.5, Mean = 84.5

4. A study was done in 2014 to investigate the relationship between playing college football and the volume (in microliters) of the hippocampus (the part of the brain that is the center of memory, emotion, and the autonomic nervous system). The study included three groups of 25 men: healthy controls who had never played football, football players with no history of concussions, and football players with a history of concussions.

(a) This study is (circle one) a controlled experiment an observational study

(b) What is the “Treatment” (explanatory) variable? _____

Is this variable categorical (qualitative) or numerical (quantitative)? CATEGORICAL NUMERICAL

(c) What is the “Response” (outcome) variable? _____

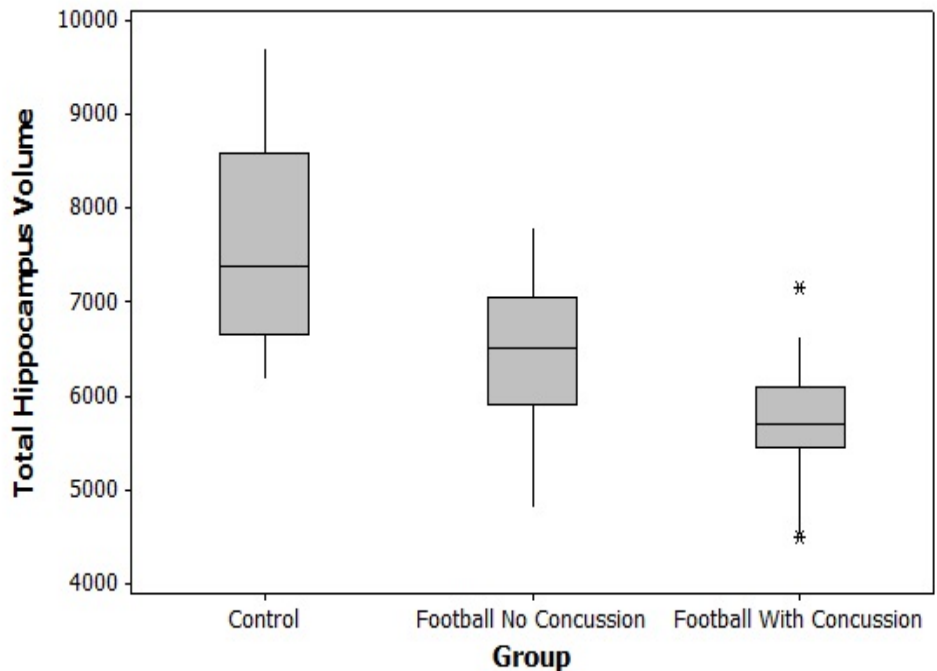
Is this variable categorical (qualitative) or numerical (quantitative)? CATEGORICAL NUMERICAL

The figure below shows side-by-side boxplots for total hippocampus volume, in microliters.

(d) Which group’s hippocampus data showed the most variability?

(e) Which group’s hippocampus data were the most symmetrically distributed?

(f) Which group had outliers?



(g) True or false: Typically, the group that didn’t play football had greater hippocampus volume than those who did play football.

(h) True or false: All of the subjects who didn’t play football had greater hippocampus volume greater than all of the subjects who did play football.

(i) True or false: There is an association between playing college football and hippocampus size.

(j) Can we conclude that playing college football causes a reduction in the size of the hippocampus? Why or why not?

5. Suppose you select a sample of five students and ask them how many texts they sent during class that day. The data values are 0, 1, 0, 12, 2
- a) Construct a dotplot for this data

 - b) Find the sample mean for the data. Plot it on the dotplot.

 - c) Find the sample median for the data. Plot it on the dotplot.

 - d) Which statistic is a more “typical” value for this data: the mean the median

 - e) Which value appears to be an outlier?

 - f) What effect did the outlier have on the mean?

 - g) Because of this effect, we say that the mean is not _____

 - h) Find the sample standard deviation by hand. Show work!

6. (9 pts) Three-year-old boys have a distribution of heights which is unimodal and symmetric, with mean of 38 inches and a standard deviation of 2 inches.

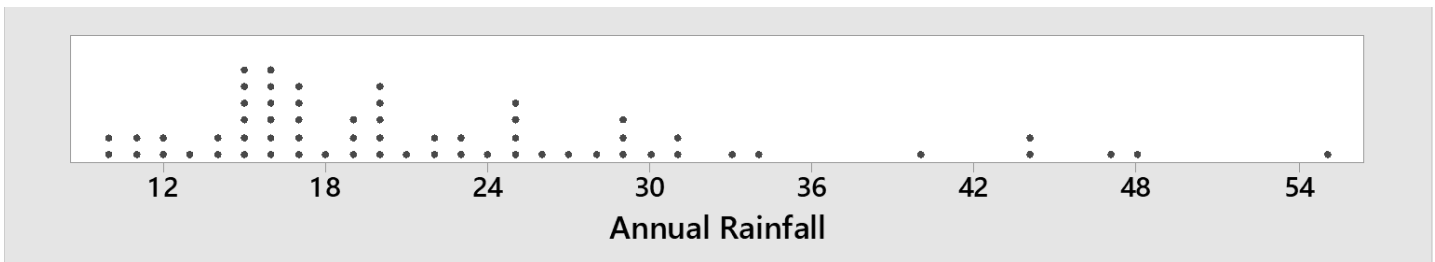
(a) Sketch a curve, with the x-axis labeled appropriately, showing the distribution of heights, with the standard deviation included as part of the sketch.

(b) Between what two values should about 95% of the heights fall? Include units in your answer.

(c) Find the z-score for a three-year-old boy's height of 43 inches.

(d) Suppose there are two children in a family, a three-year-old boy who is 43 inches tall and a ten-year-old girl who is 61 inches tall. If the mean and standard deviation for the heights of ten-year-old girls is 54.5 inches and 2.5 inches, determine which child is unusually tall for his/her age.

7. The annual rainfall (in inches) for a number of years in San Luis Obispo is graphed below, with the summary statistics from Minitab shown below that. Use the graph and the summary stats to answer the questions.



Descriptive Statistics: Annual Rainfall

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Annual Rainfall	62	0	22.47	1.27	9.97	10.35	15.52	19.73	27.36	54.53

- (a) How many data values are there?
- (b) What percentage of the years had an annual rainfall total that was over 36 inches?
- (c) Which would it be more appropriate to describe the center and variation of this data set: (circle one)
 - the mean and standard deviation
 - the median and IQR
- (d) What is the five number summary for this data set?
- (e) Find the IQR.
- (f) Find the Lower Outlier and Upper Outlier Limits.
- (g) What values are outliers? Include units!

Math 247: Test 1 (Take Home)

Name: _____

(20 points)

In-Class Test _____/80

Take-Home Test _____/20

- This exam is due at the beginning of class on Thursday, 2/8/2018. You may work with other people in the class but not with tutors, other instructors, etc. Be sure that all answers are written in your own words; i.e., do not write verbatim the same answer as another student.
- Your work can be typed or **NEATLY** handwritten.
- Your work should incorporate all of the technology work mentioned below; i.e., **copy and paste the Minitab work into your write-up.**

Scoring will be based on organization of your work, accuracy, and thoughtful, well-written answers.

Answer all of the following questions on another piece of paper and attach to this cover sheet. Use complete sentences in your answers!

Pandora took a random sample of 25 listeners to find out how many minutes they spent listening to internet radio during the last week. The (sorted) results were as follows:

0	170	180	200	240	230
230	250	280	280	290	310
310	310	330	350	350	360
370	390	400	400	430	450
480					

1.) Use Minitab to graph a histogram from the data. Adjust the classes (bins) so that there are 9 classes (intervals) and they begin at 0. You'll need to double click on the bars of the histogram, select the "binning" tab, then select "cutpoint" and "number of intervals".

- Describe the distribution, including shape, spread, potential outliers, and "typical value".

2.) Use Minitab to find the Descriptive Statistics" for the data.

- Explain whether or not it be appropriate to apply the Empirical Rule to this data set. Write a sentence or two (complete sentences, please!) to explain and reference the histogram in your explanation.
- By hand, showing work, find the range of listening times that are 1 standard deviation from the mean, 2 standard deviations from the mean, 3 standard deviations from the mean:

Fill these values by hand on your histogram.

- Determine how many actual data values (using the given Pandora data) fall in each of the ranges above, then compute what percent of the total (total of 25 data values) is in each range: Report these numbers and percentages as shown below:

Actual data within 1 standard deviation of the mean: _____/25 = _____ %

Actual data within 2 standard deviations of the mean: _____/25 = _____ %

Actual data within 3 standard deviations of the mean: _____/25 = _____ %

- Relate the actual data percentages to those predicted by the Empirical Rule? Include the predicted values from the Empirical Rule as well as the percentages you found in part (e).

3.) By hand, find the outlier limits (show work!) and explain what this tells you about any potential outliers.

- Use Minitab to graph a boxplot and explain whether or not it confirms your outlier analysis.
- Find the z-score for the outlier and interpret it in the context of the problem. In your explanation, connect the z-score to what it means to be an outlier.