DHS AP Statistics Summer Packet

- In this class you will learn to describe and analyze sets of data and use that analysis to draw conclusions in context about the situation that gave the data. This course is not like any other math class you have ever taken! I would say that it is a combination of Math, English, and Science. Communication skills are essential, and there is much more reading and writing than what you are used to in a math class. It is a very rewarding course and a very important one, in my opinion, but can be quite difficult at times. Since it is an AP course, it is considered to be college-level. The mathematics required for this course may not be as difficult as in other advanced math courses, but some of the concepts can be very confusing. In addition, there is a great deal of material that we will cover by the end of the semester, so the class moves fast.

- The summer assignment is designed for you to review some algebra skills, general information, and refresh your knowledge or learn about basic descriptive statistics, graphs, etc.. Answers should be neatly handwritten. You must have a graphing calculator for this class (TI-84 is recommended). Calculate84 is also a FREE app that emulates the TI-84 calculator on your phone screen. You can use the reference pages at the beginning of the packet and/or watch YouTube videos for calculator tips.

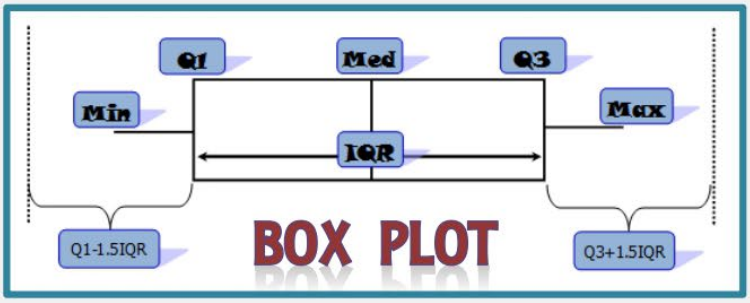
- Due Date: Summer assignments must be completed by the first week of class. The class will NOT go slowly to accommodate those who did not complete the summer work.

- Supplies needed for this course are the following;
    - 3-ring binder with loose leaf paper and dividers (5)
    - Pencils and erasers
    - Highlighters, colored pens, colored pencils, etc.
    - Calculator: TI-83 or TI-84 graphing calculator

- What You Can Expect:
    - Direct instruction
    - REGULAR HOMEWORK
    - Show all of your work like teacher does in class
    - Explaining in words how your results relate to the problem given

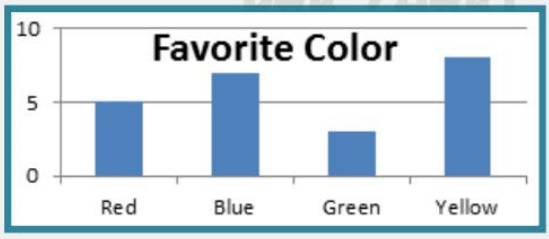
- Ways to Succeed in this Class
    - Take notes <u>every day</u> and keep them organized by chapter/date in your binder
    - Fully complete homework every day and check your solutions
    - Participate in class
    - Read your text book

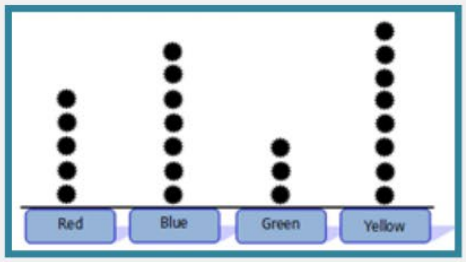
Graphical Displays



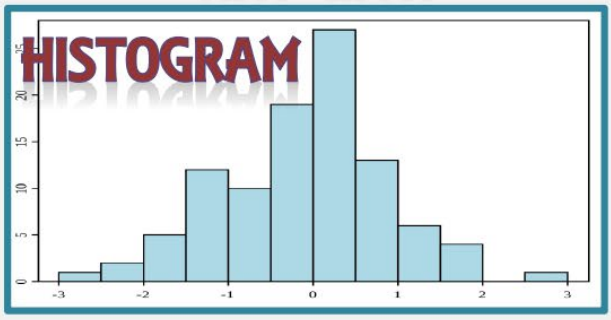
BOX PLOT

BAR CHART





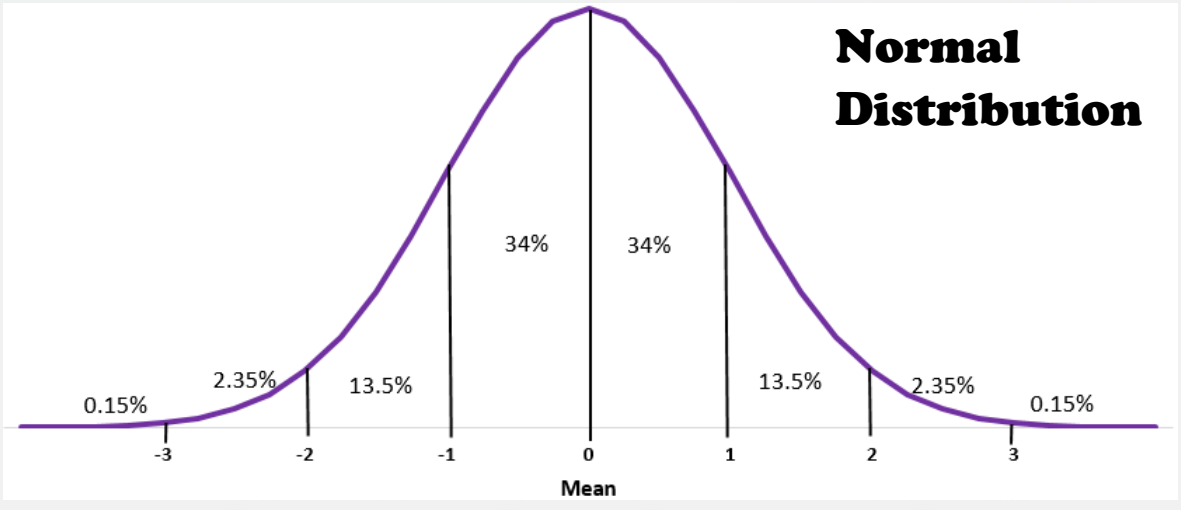
DOT PLOT



Per Capita GNP
Western Africa

Stem	Leaf
1	80
2	40 60 70
3	10 30 60 70 90
4	10 80
5	00
6	
7	10 30
8	90

STEM & LEAF



Frequently Used Formulas

Z-Statistic

1-Proportion z-interval $\hat{p} \pm z^* \sqrt{\frac{\hat{p}\hat{q}}{n}}$

1-Proportion z-test $z = \frac{(\hat{p}-\rho)}{\sqrt{\frac{\rho q}{n}}}$

2-Proportion z-interval $\hat{p}_1 - \hat{p}_2 \pm z^* \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$

2-Proportion z-test $z = \frac{(\hat{p}_1-\hat{p}_2)-0}{SE_{pooled}(\hat{p}_1-\hat{p}_2)}$

$SE_{pooled}(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{\hat{p}_{pooled}\hat{q}_{pooled}}{n_1} + \frac{\hat{p}_{pooled}\hat{q}_{pooled}}{n_2}}$

T-Statistic

1-Sample t-interval $\bar{x} \pm t^*_{n-1}\left(\frac{s}{\sqrt{n}}\right)$

1-Sample t-test $t_{n-1} = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}}$

2-Sampe t-interval $(\bar{x}_1 - \bar{x}_2) \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

2-Sample t-test $t_{n-1} = \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$

Matched Pairs t-test $t_{n-1} = \frac{\bar{d}-\mu_d}{\frac{s_d}{\sqrt{n}}}$ where d represents the difference between pairs of data

χ^2 statistic

$\sum \frac{(Observed-Expected)^2}{Expected}$

Common Calculator Commands

5-Number Summary

STAT-EDIT-ENTER DATA
STAT-CALC-1-Var Stats

Box Plot

STAT-EDIT-ENTER DATA
2ND-Y=-TURN PLOT ON
TYPE: CHOOSE BOXPLOT W/DOTS
CHOOSE Xlist
ZOOM-9:ZoomStat

Histogram

STAT-EDIT-ENTER DATA
2nd-Y=-TURN PLOT ON
TYPE: CHOOSE HISTOGRAM
CHOOSE Xlist
ZOOM-9:ZoomStat

Diagnostics On

2nd-0-x^{-1} – DiagnosticOn-ENTER

Statistics Tests

2nd – VARS
STAT – CALC
STAT - TESTS

DHS AP Statistics Summer Packet

Part 1: Vocabulary. Please define each of the following terms from the information on www.stattrek.com. When asked to provide and example of the word, provide a unique example of the word NOT given on the website.

1. Categorical Variables

Example:

2. Quantitative Variables

Example:

3. Univariate Data

Example:

4. Bivariate Data

Example:

5. Sample

Example:

6. Symmetry

Example:

7. Skewness

Sketch Skewed left:

Sketch Skewed right:

Part 2: Algebra review. While there isn't a great deal of algebra and equation solving in statistics, some parts of algebra 1 and algebra 2 will come in handy. Please complete the following, showing your work for each part.

DHS AP Statistics Summer Packet

Here is a formula that is used often in AP Statistics: $z = \frac{x - \bar{x}}{s}$

1. If z = 2.5, x = 102, and $\bar{x}$=100. What is s? Show your work.

2. If z = -3.35, x = 60, and s = 4, what is $\bar{x}$? Show your work.

3. Solve $0.05 = 1.96\sqrt{\frac{0.5^2}{n}}$ for n.

4. If $-1.64 = \frac{60 - \mu}{\sigma}$ and $1.96 = \frac{95 - \mu}{\sigma}$, solve for μ and σ.

Evaluate Expressions:

1. $\frac{x-\bar{x}}{s}$ *when* $x = 83, \bar{x} = 91$ *and* $s = 14$

2. $\sqrt{\frac{pq}{n}}$ *when* $p = .30, q = .70$ *and* $n = 23$

3. $\frac{s}{\sqrt{n}}$ *when* $s = 17.03$ *and* $n = 20$

4. $\hat{y} = 63.1 - 12.3x$ *find* $\hat{y}$ *when* $x = 4$

5. $\sqrt{\frac{p(1-p)}{n_1} + \frac{p(1-p)}{n_2}}$ *when* $p = .34, n_1 = 24$ *and* $n_2 = 31$

Part 3: Calculator Practice. Use the steps below to familiarize yourself with the basic display functions of your calculator.

DHS AP Statistics Summer Packet

Histograms: A set of quantitative data is often separated into groups or intervals. For example, test scores may be separated into 90-99, 80-89, 70-79, and so on. The frequency refers to the count for each interval. A histogram is a visual representation of the frequencies plotted against the interval.

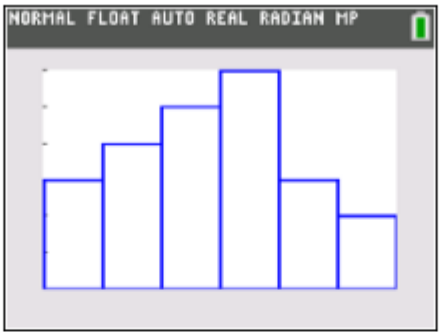
Exercise 1: Enter the following test scores into List 3 and create a histogram using intervals of width 10:

100, 100, 98, 95, 92, 88, 85, 84, 84, 81, 80, 77, 75, 72, 70, 70, 66, 65, 63, 60, 59, 55, 50

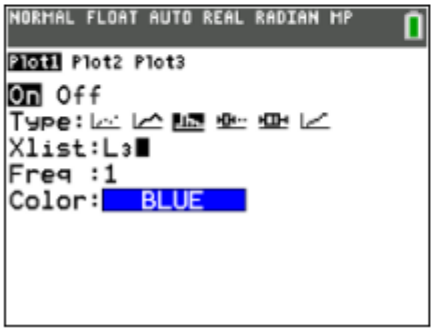
Procedure: First enter the scores into L3. You must now turn on your STATPLOT. The 2nd function of the Y= will get you to the menu.

Choose one of the Plots by pressing 1, 2, or 3. Now activate the plot by selecting ON. At the TYPE prompt choose the third icon for histogram. At the X LIST, make sure it says L3 (2nd function on the numbers 1 to 6). Frequency should be 1.

To control the width of the bars (called the bin width), set up the window manually. Xscl=10



You may now TRACE the graph to see the intervals and their frequencies. When drawing a histogram, make sure both axes are labeled and have appropriate scales.



Now go back to the WINDOW for this example and change the Xscl to 5 and GRAPH. Draw the histogram below and describe how this one differs from the first histogram. Which histogram would best describe the data?

Go back and change the Xscl to 15 then graph. Describe how this histogram differs from the first two. Aagain, state which histogram would best describe the data.

From these two problems, what conclusion can you draw about how the Xscl affects the histogram?

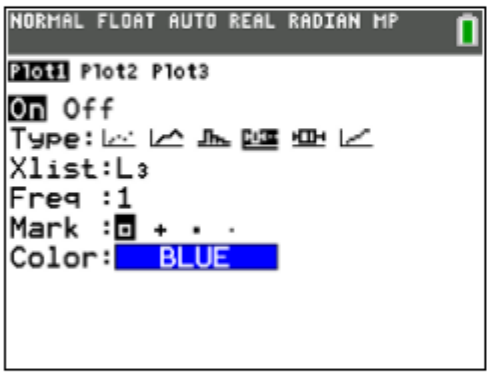
DHS AP Statistics Summer Packet

A box plot is a graphical display for a set of quantitative data that only displays the five number summary:

minimum, Q1, median, Q3, maximum

Using the same data as before (example 1 - test scores), create a boxplot below. Make sure your axis has a scale and label.

In the STATPLOT menu, you will notice two boxplot options. The first one allows for possible outliers. I recommend you always use this one.



Suppose another class receives these test scores:

94, 93, 90, 83, 82, 82, 77, 76, 75, 70, 64, 60, 53, 52, 48, 44, 40

Create two box and whiskers plots on the same set of axes to compare their performance with the original group.

Write a few sentences comparing the medians and the spreads of each class.

DHS AP Statistics Summer Packet

Part 4: Probability. Using the information from the 2-way table presented below, answer the questions. Express all answers as reduced fractions and then decimals rounded to the nearest ten thousandth (4 decimal places).

A survey was done of commuters in three major cities about how they primarily got to work. The results are shown in the frequency table below.

	Car	Train	Walk	Total
New York	10	50	20	80
Los Angeles	36	24	10	70
Chicago	16	28	6	50
Total	62	102	36	200

Find the following probabilities if one commuter is selected at random.

a.) P(NY)

b.) P(NY and Car)

c.) P(NY or Car)

d.) P(Train | NY)

e.) P(Train | LA)

f.) P(Train | Chicago)

g.) In which city is a commuter most likely to take the train as a method of transportation? Explain your answer.

DHS AP Statistics Summer Packet

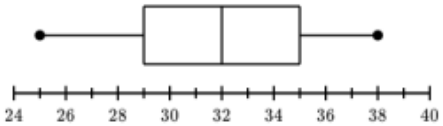
Part 5: Shapes of graphs. Using the information provided, analyze the graphical displays and answer the questions.

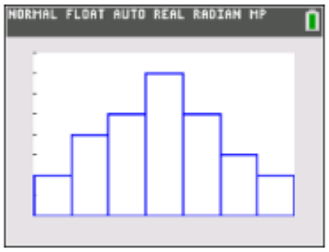
When describing a set of data we look at the following features:

Shape center spread clusters and gaps outliers

We have several terms that we use to describe the shape but this packet will concentrate on only two: symmetric and skewed.

You can tell if a graph is symmetric if a vertical line in the center divides the graph into two fairly congruent shapes. The following sets of data can be described as symmetric:

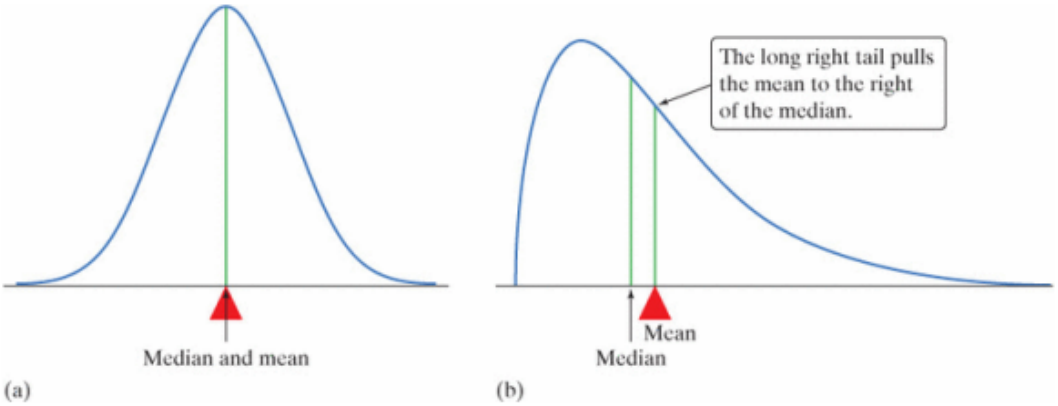




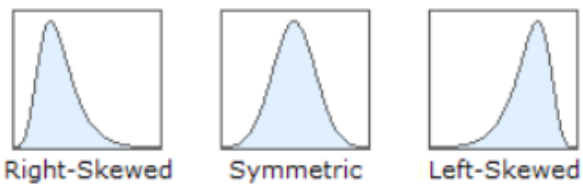
The mean and the median are approximately the same in a symmetric graph. Figure (a) below.

One can tell if a graph is skewed if the graph has a big clump of data on either the left or the right with a tendency to get flatter on the opposite side.

One can tell if a graph is skewed if the graph has a big clump of data on either the left or the right with a tendency to get flatter on the opposite side.



The distribution in figure (b) is skewed right. The mean gets pulled toward the tail, which is the direction of the skewness. So the mean > median

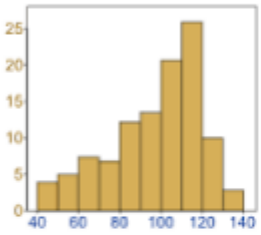


The mean is always pulled closer toward the tail.

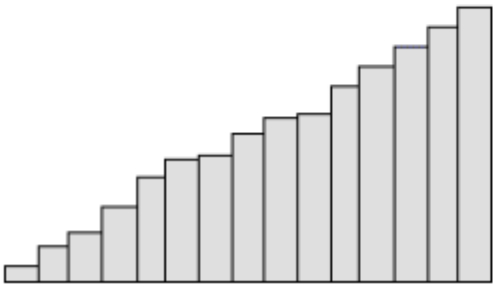
DHS AP Statistics Summer Packet

Exercise 1. Determine if the distribution is symmetric, skewed left or skewed right. Then determine if the mean or the median is the larger value. If they are equal, state so.

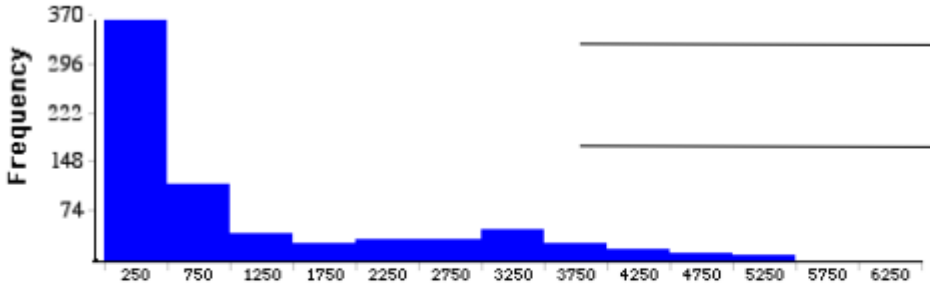
1.



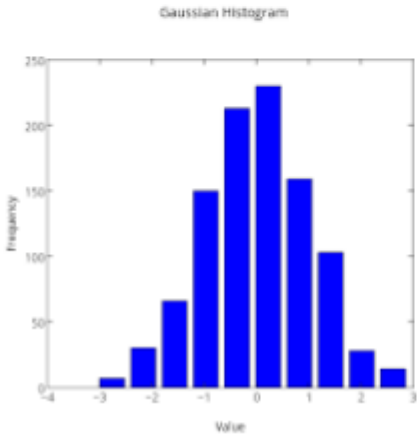
2.



3.

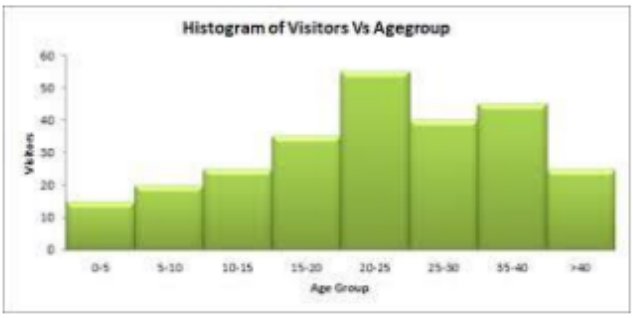


4.

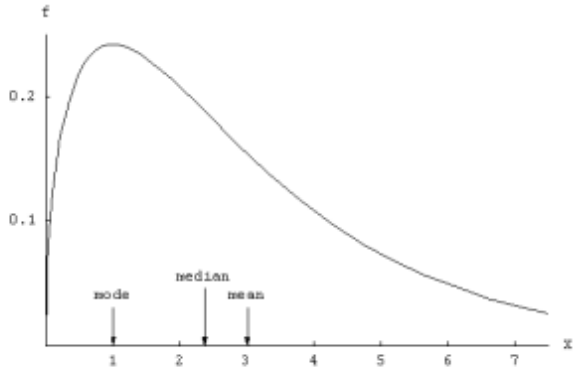


DHS AP Statistics Summer Packet

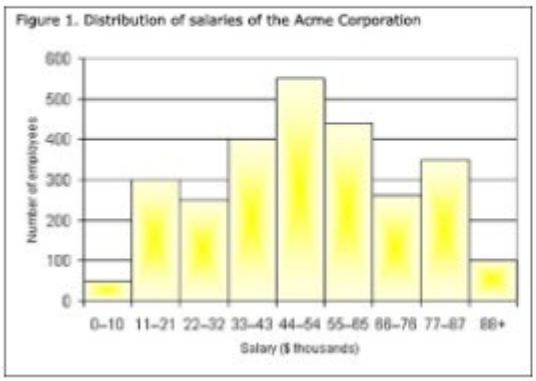
5.



6.



7.

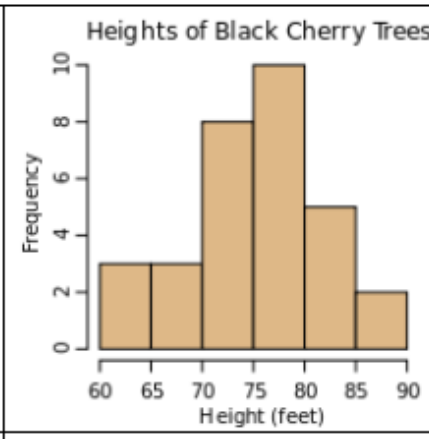


DHS AP Statistics Summer Packet

8. For the following graphs, find the shape, center (use the median), and spread (use the range). If there are any other notable features evident in the graph (clusters, gaps or outliers), then say where they are. Otherwise, do not comment on them.

Note: To find the center of these graphs, use the frequencies found on the y-axis. Count how many are in each bar. If the bar has an interval, use the midpoint. Add these up and divide by 2. This tells you where the median is located. Which bar is this value located in?

1



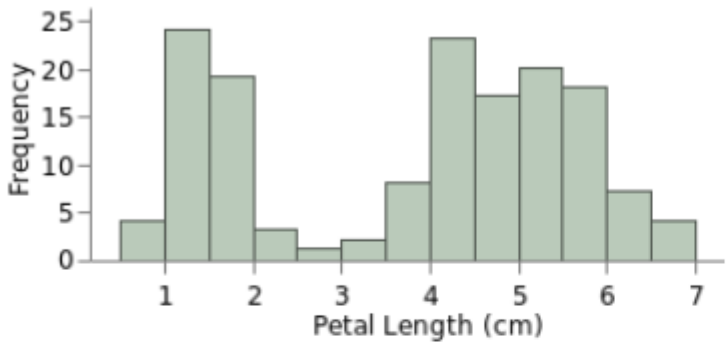
Shape _____

Center _____

Spread _____

Clusters? _____

2



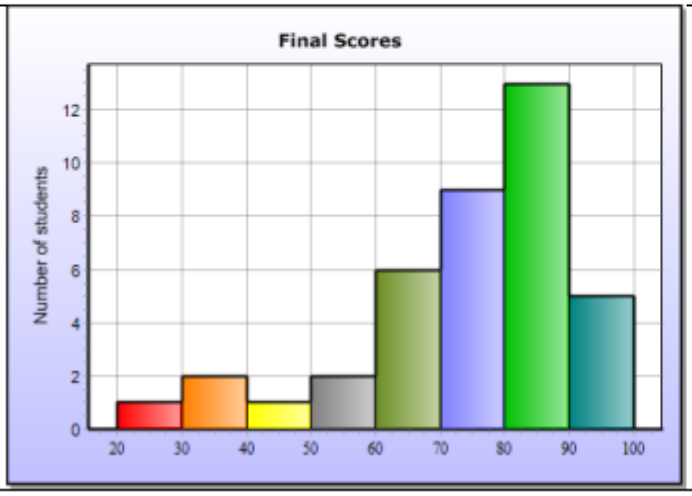
Shape _____

Center _____

Spread _____

Clusters? _____

3



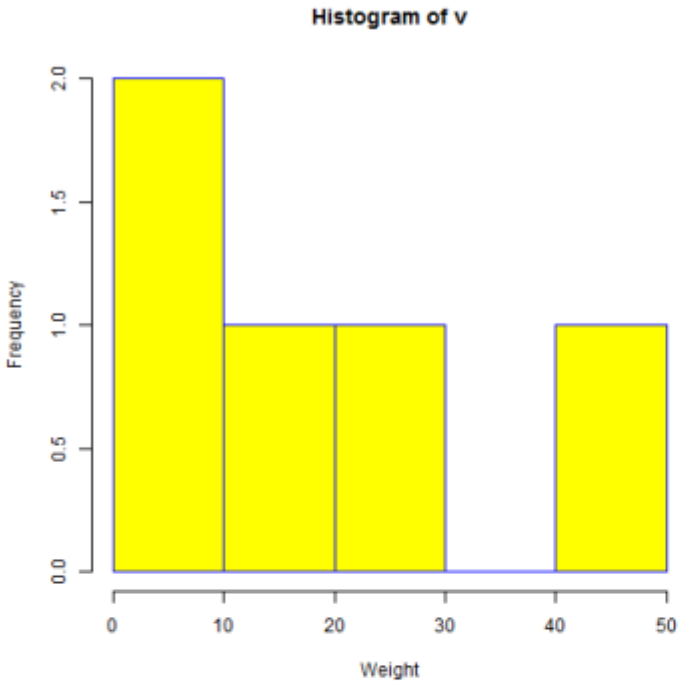
Shape _____

Center _____

Spread _____

Clusters? _____

4



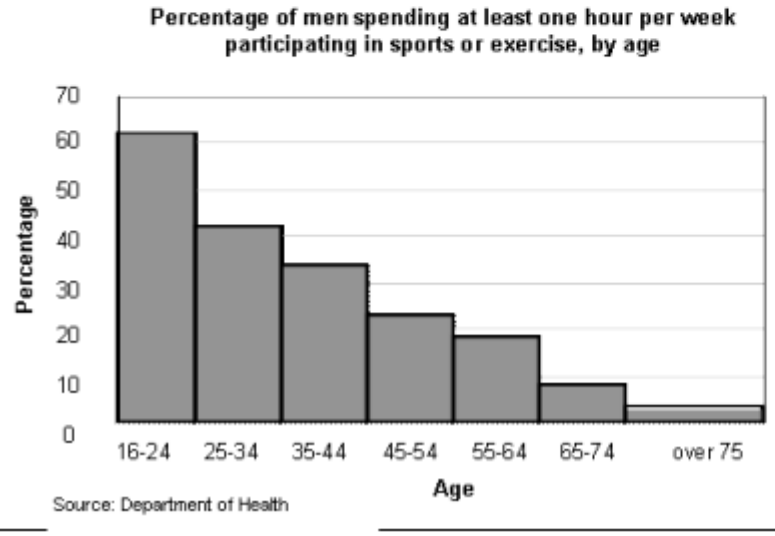
Shape _____

Center _____

Spread _____

Clusters? _____

5



Shape _____

Center _____

Spread _____

Clusters? _____

Thank you to wdupuis of Helena HS.