Unit 4: worksheet #1

Name Key

Hour_____

1. Here is a portion of Tampa Bay's team roster from 2003:

No.	Name	Pos.	Ht.	Wt.	Birth date	Yrs. Of Exp.	College
14	B. Johnson	QB	77	224	09/13/68	9	Florida State
99	W. Sapp	DT	74	303	12/06/72	8	Miami
7	M. Gramatica	K	68	170	11/27/75	4	Kansas State
9	T. Tupa	P	76	235	02/06/66	14	Ohio

(a) What are the qualitative variables?

Name, Pos, Birth date, College, No.

(b) What are the quantitative variables?

Ht., Wt., Yrs. of Exp.

2. A contest is held to determine the household that recycles the most. Weight may not be the best way to measure the participation of households the recycling program. What variables could you use to measure the level of participation instead?

3. Colleges and universities are often ranked on their "academic quality." Describe <u>two qualitative</u> and <u>two quantitative</u> variables that you would like to know when you are trying to pick a college to attend.

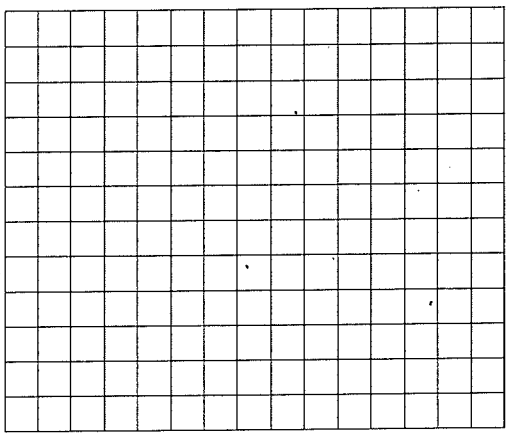
quantitative: population, class sizes

qualitative: majors offered, location

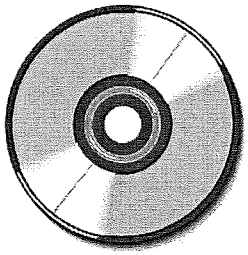
4. A survey of college freshmen in 2001 asked what field they planned to study. The results: 12.6% arts and humanities, 16.6% business, 10.1% education, 16% engineering and science, 12% professional, and 10.3% social science.

(a) What percent plan to study fields other than those listed?

(b) Make a bar graph that displays the data.

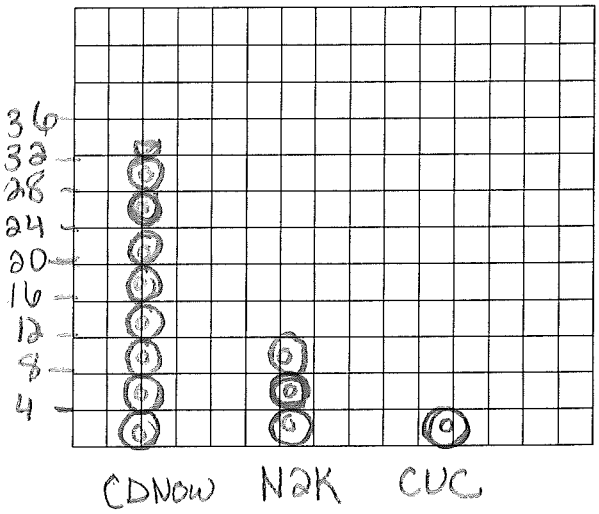


5. This pictogram shows the market leaders for online CD sales in 1997.
Make an <u>improved</u> pictogram that more fairly and accurately portrays the data.



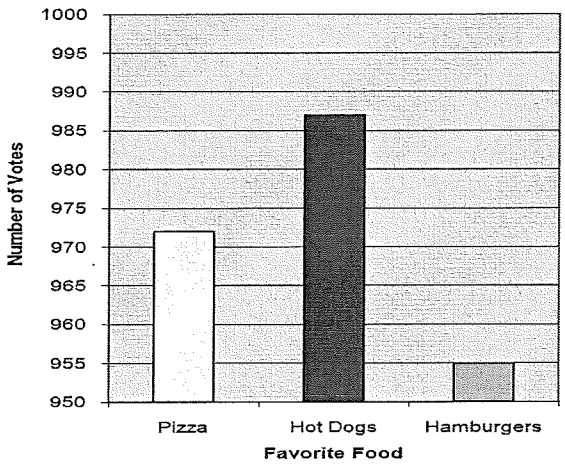
CDNow 33% **N2K 12%** **CUC 4%**

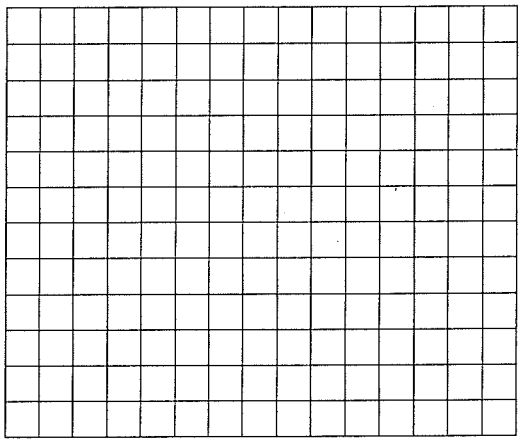
Online CD Sales Leaders, 1997



Online CD Sales Leaders, 1997

6. This bar graph shows the results of a survey taken at a Twins game in 2008.
Make an improved bar graph that gives a more fair comparison between the favorites.





Unit 4: worksheet #2

Name Key

Hour

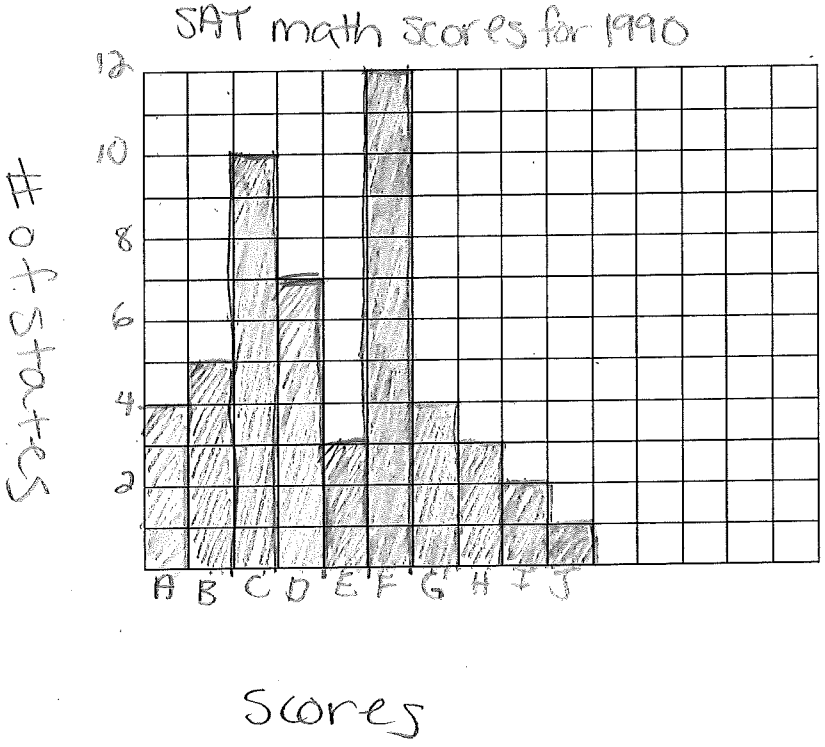
1. This table shows average SAT math scores for 1990.

State	Score	State	Score	State	Score	State	Score
Alabama	514	Illinois	528	Montana	523	Rhode Island	461
Alaska	476	Indiana	459	Nebraska	546	South Carolina	487
Arizona	497	Iowa	577	Nevada	487	South Dakota	555
Arkansas	511	Kansas	548	New Hampshire	486	Tennessee	525
California	484	Kentucky	521	New Jersey	473	Texas	461
Colorado	513	Louisiana	517	New Mexico	527	Utah	539
Connecticut	471	Maine	463	New York	470	Vermont	468
Delaware	470	Maryland	478	North Carolina	440	Virginia	470
DC	441	Massachusetts	473	North Dakota	564	Washington	486
Florida	466	Michigan	514	Ohio	499	West Virginia	490
Georgia	443	Minnesota	542	Oklahoma	523	Wisconsin	523
Hawaii	481	Mississippi	519	Oregon	484	Wyoming	519
Idaho	502	Missouri	522	Pennsylvania	463		

(a) Make a frequency table and a histogram for these data. Use a class width of 15.

min: 437 max: 577

Classes	frequency	
135–149 A	1111	40
150–164 B	1111	5
165–179 C	1111 1111	10
180–194 D	1111 11	7
195–209 E	111	3
210–224 F	1111 1111 11	12
225–239 G	1111	4
240–254 H	111	3
255–269 I	11	2
270–284 J	1	1



(b) Describe the distribution. (Remember your SOCS!)

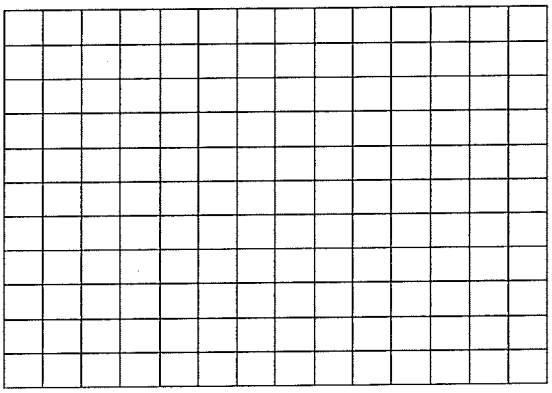
S: Symmetric
O: None
C: mean = 497
S: 437 – 577

(c) Compare the average SAT scores for coastal states to those of non-coastal states. Identify the connection between the geographical location and average SAT math score.

2. This table shows the percent of people in each state who are at least 65 years old.

AL	11.5	HI	8.2	MA	12.8	NM	9.1	SD	13.5
AK	3.0	ID	10.2	MI	10.2	NY	12.4	TN	11.5
AZ	11.7	IL	11.2	MN	12.0	NC	10.6	TX	9.6
AR	13.9	IN	11.1	MS	11.7	ND	12.4	UT	7.5
CA	10.3	IA	12.0	MO	13.3	OH	11.1	VT	11.5
CO	8.6	KS	13.1	MT	11.0	OK	12.3	VA	9.6
CT	12.0	KY	11.4	NE	13.2	OR	11.9	WA	10.6
DE	10.3	LA	9.6	NV	8.5	PA	13.2	WV	12.4
FL	17.3	ME	12.7	NH	12.7	RI	13.6	WI	12.2
GA	9.6	MD	9.6	NJ	11.9	SC	9.5	WY	7.8

(a) Make a frequency table and a histogram for these data. Use a class width of 2.



(b) Describe the distribution.
(Remember your SOCS!)

3. Here are results of a survey of 7th grade girls.

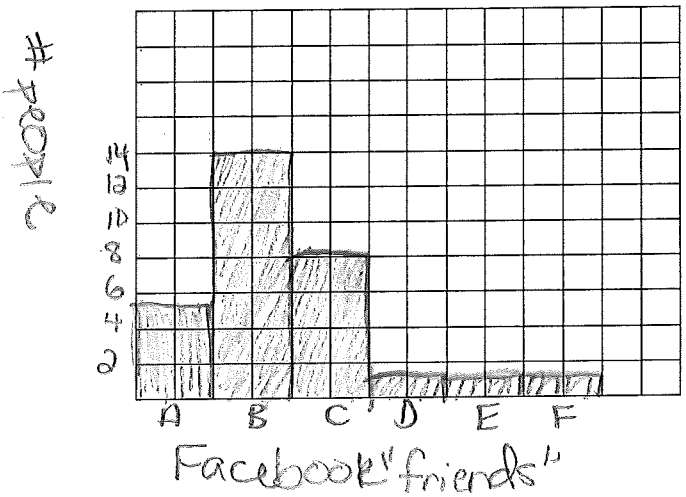
Number of Facebook "friends"

19, 25, 33, 26, 38, 11, 53, 42, 39, 21,
99, 32, 75, 18, 27, 35, 40, 27, 36, 50,
41, 33, 25, 60, 44, 35, 30, 24, 52, 38

Min: 11
Max: 99

(a) Make a frequency table and a histogram for these data. Use a class width of 15.

Classes		frequency
10-24	A	5
25-39	B	14
40-54	C	8
55-69	D	1
70-84	E	1
85-99	F	1



(b) Describe the distribution.
(Remember your SOCS!)

S: Skewed right
O: None
C: mean = 37.6
S: 11 - 99

Unit 4: worksheet #3

Name Key

Hour_____

1. Here are the number of home runs that Babe Ruth hit in each of his 15 years with the New York Yankees, 1920 - 1934: 54 59 35 41 46 25 47 60 54 46 49 46 41 34 22

(a) Make a stemplot for the data.

2	25
3	45
4	1166679
5	449
6	0

(b) Describe the distribution (SOCS).

S: 22 - 60
O: None
C: Med = 46
S: Symmetric

2. Here are the average monthly rainfall amounts for Chicago and Miami in 2009.
Chicago rainfall: 30", 26", 37", 16", 19", 24", 29", 40", 36", 27", 36", 36"
Miami rainfall: 19", 20", 31", 59", 70", 90", 85", 60", 25", 32", 21", 78"

(a) Make a back-to-back stemplot comparing the data sets.

(b) Describe the Chicago distribution.

(c) Describe the Miami distribution.

(d) What are the main differences you see between the data sets?

3. Nolan Ryan holds the record for most career strikeouts by a major league pitcher. Ryan's season strikeout totals were: 6, 133, 92, 125, 137, 329, 383, 367, 186, 327, 341, 260, 223, 200, 140, 245, 183, 197, 209, 194, 270, 228, 301, 232, 203, 157, 46

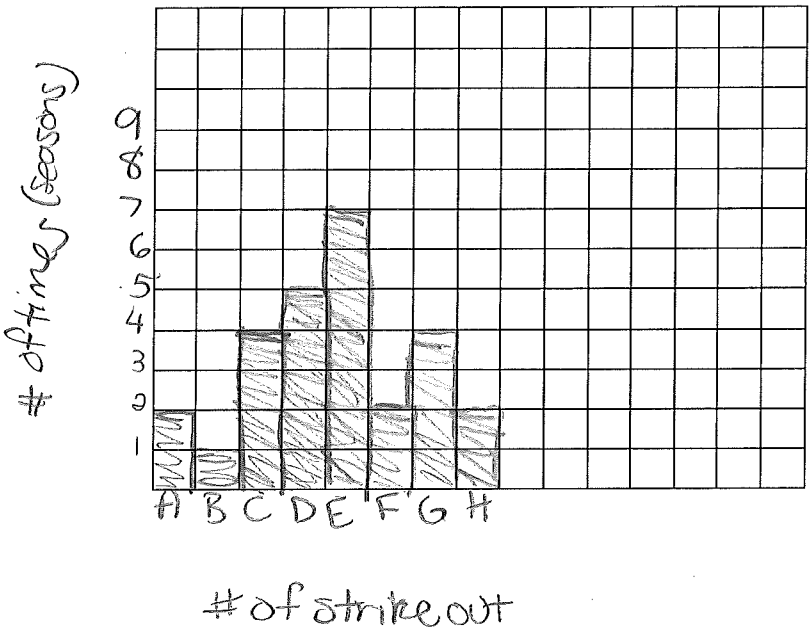
(a) A stemplot would not be the best choice for this data set. Why?

The numbers are too spread out

(b) Make a frequency table and a histogram for these data. Use a class width of 50.

min = 6 max = 383 strikeouts (Nolan Ryan)

	Classes	frequency	
A	00-49	II	2
B	50-99	I	1
C	100-149	IIII	4
D	150-199	HHT	5
E	200-249	HHTII	7
F	250-299	II	2
G	300-349	IIII	4
H	350-399	II	2



(c) Describe the distribution.

S: 6-383

O: None

C: mean = 208

S: Symmetric

4. What are the advantages of histograms over stemplots?

It is an easy way to display a lot of data or spread out data.

5. What are the advantages of stemplots over histograms?

It displays your data points.

Unit 4: worksheet #4

Name Key

Hour_____

1. *Consumer Reports* magazine presented the following data on the number of calories in a hotdog for different17 brands.

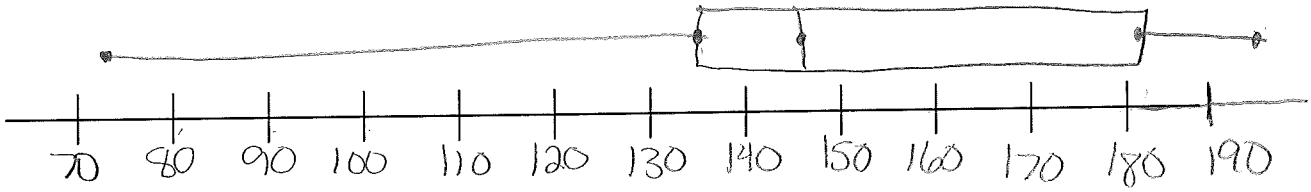
73 191 182 190 172 147 146 139 175 136 179 153 107 195 135 140 138

(a) Find the 5 number summary.

73, 107, 135, 136, 138, 139, 140, 146, 147, 153, 172, 175, 179, 182, 190, 191, 195

Min = 73
Q_1 = 137
Med = 147
Q_3 = 180.5
Max = 195

(b) Make a boxplot for the hotdog data.



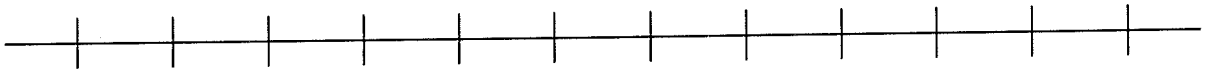
Calories in a hotdog

2. The 27 students in a statistics class kept a record of the number of hours they spent on homework in one week. These are the results:

9.5, 14.0, 7.0, 7.5, 12.0, 8.5, 11.5, 10.0, 7.0, 16.0, 10.0, 11.0, 9.0, 11.5,
8.0, 14.0, 10.5, 12.5, 7.0, 13.0, 13.5, 10.0, 6.5, 7.0, 11.0, 9.0, 6.0

(a) Find the 5 number summary.

(b) Make a boxplot for the homework data.



3. Barry Bonds and Hank Aaron are two of the most prolific homerun hitters in the history of major league baseball.

Bonds # of HRs/season: 16 25 24 19 33 25 34 46 37 33 42 40 37 34 49 73 46

16 19 24 25 25 33 33 34 37 37 40 42 44 46 46 49 73

Aaron # of HRs/season: 13 20 24 26 27 29 30 32 34 34 38 39 40 40 44 44 44 44 45 47

28 36

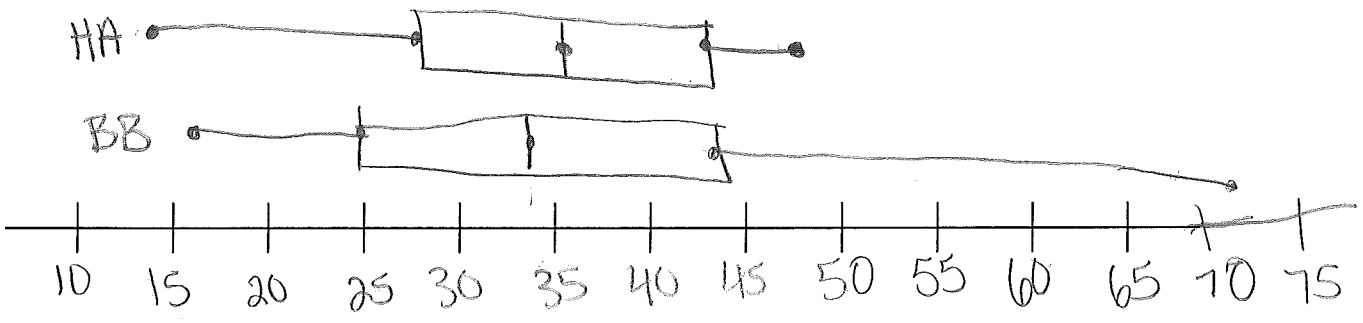
(a) Find the 5 number summary for Barry Bonds' data.

Min = 16
Q_1 = 25
Med = 34
Q_3 = 44
Max = 73

(b) Find the 5 number summary for Hank Aaron's data.

Min = 13
Q_1 = 28
Med = 36
Q_3 = 44
Max = 47

(c) Make a "stacked" boxplot of both sets of data on one plot.



(d) Describe the Barry Bond distribution. (show calculation for outliers)

S: 16 to 73
O: yes, 73
C: Med = 34
S: skewed right

$Q_1 - 1.5(IQR)$
$25 - 1.5(19) = -3.5$
$Q_3 + 1.5(IQR)$
$44 + 1.5(19) = 72.5$

(e) Describe the Hank Aaron distribution. (show calculation for outliers)

S: 13 to 47
O: None
C: med = 36
S: skewed left

$Q_1 - 1.5(IQR)$
$28 - 1.5(16) = 4$
$Q_3 + 1.5(IQR)$
$44 + 24 = 68$

Unit 4: ws #4B Name Key

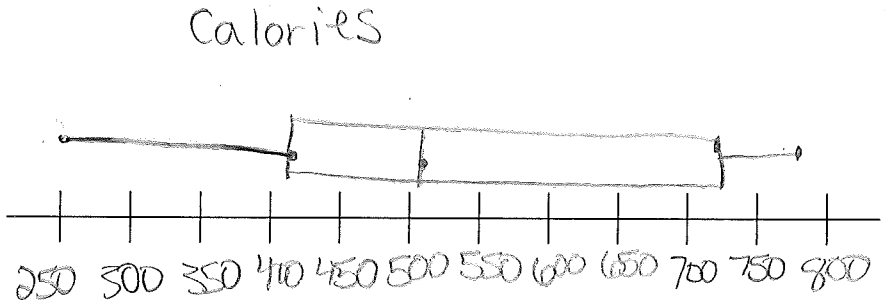
1. Use the data in the table to complete the following questions.

Nutrition Facts	Serving Size	Calories	Calories from Fat	Total Fat (g)
Sandwiches				
Hamburger	3.5 oz (100 g)	250	80	9
Cheeseburger	4 oz (114 g)	300	110	12
Double Cheeseburger	5.8 oz (165 g)	440	210	23
McDouble	5.3 oz (151 g)	390	170	19
Quarter Pounder® with Cheese+	7 oz (198 g)	510	230	26
Double Quarter Pounder® with Cheese++	9.8 oz (279 g)	740	380	42
Big Mac®	7.5 oz (214 g)	540	260	29
Big N' Tasty®	7.2 oz (206 g)	460	220	24
Big N' Tasty® with Cheese	7.7 oz (220 g)	510	250	28
Angus Bacon & Cheese	10.2 oz (291 g)	790	350	39
Angus Deluxe	11.1 oz (314 g)	750	350	39
Angus Mushroom & Swiss	10 oz (283 g)	770	360	40

a) Calculate the 5 number summary for the "calories" data.

Min = 250
Q_1 = 415
Med = 510
Q_3 = 745
max = 790

b) Construct a boxplot for the "calories" data.



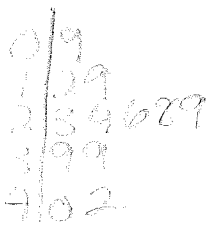
c) Describe the "calories" distribution. (Make sure to use the formula to determine if there are outliers.)

S: 250 to 790
O: None
C: Med = 510
S: skewed left.

$Q_1 - 1.5(IQR)$
$415 - 1.5(330) = -80$
$Q_3 + 1.5(IQR)$
$745 + 1.5(330) = 1240$

d) Make a stemplot using the "total fat" data.

9 12 19 23 24 26 28 29 37 39 40 42



e) Describe the "total fat" distribution. (Make sure to use the formula to determine if there are outliers.)

S: 9 to 42
O: None
C: med = 27
S: Symmetric

$Q_1 - 1.5(IQR) = 21 - 1.5(18) = -6$
$Q_3 + 1.5(IQR) = 39 + 1.5(18) = 66$

2. Use the given stemplot to complete the following questions.

a) Calculate the 5 number summary.

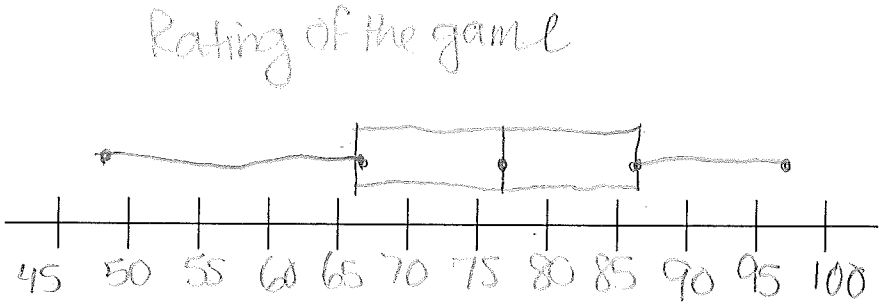
min = 48
Q_1 = 66.5
med = 77
Q_3 = 86
max = 97

Ratings of the "Wow So Fit" Game

4	8
5	2 3 5 7
6	2 2 4 6 7 7 9
7	0 0 2 5 6 7 7 9 9
8	1 1 2 2 3 5 6 6 7 9
9	0 2 2 2 6 7

Key: 6 | 7 = 67

b) Construct a boxplot.



c) Describe the distribution. (Make sure to use the formula to determine if there are outliers.)

S: 48 to 97
O: None
C: Med = 77
S: Symmetric

$Q_1 - 1.5(IQR)$
$66.5 - 29.25 = 37.25$

$Q_3 + 1.5(IQR)$
$86 + 29.25 = 115.25$

Unit 4: worksheet #5

Name Key
Hour_____

1. A. Find the mean and standard deviation of the data set: **15, 18, 23, 29, 35**

Mean: 24
St. Dev: 8.12

B. When you add 5 to each number in the set you get: **20, 23, 28, 34, 40.**
Find the new mean and standard deviation.

Mean: 29
St. Dev: 8.12

C. When you subtract 7 to each number in the set you get: **8, 11, 16, 22, 28.**
Find the new mean and standard deviation.

Mean: 17
St. Dev: 8.12

D. Study your results in part (A) through part (C). Explain how the <u>mean</u> changes when you add or subtract a number from each member of the data set.

goes up/down by what you add/subtract

E. Explain how the <u>standard deviation</u> changes when you add or subtract a number from each member of the data set.

It stays the same

2. A. When you multiply each number in the set by 3 you get: **45, 54, 69, 87, 105.** Find the new mean and standard deviation.

Mean: 72
St. Dev: 24.4

B. When you multiply each number in the set by ½ you get: **7.5, 9, 11.5, 14.5, 17.5.** Find the new mean and standard deviation.

Mean: 12
St. Dev: 4.1

D. Compare your answers to the mean and standard deviation in #1A. Explain how the <u>mean</u> changes when you multiply each member of the data set by a number.

It is multiplied by the same number

E. Explain how the <u>standard deviation</u> changes when you multiply each member of the data set by a number.

It is multiplied by the same number

Unit 4 ws #6

Name Key

Hour _____

1. Use the data collected (number of jumping jacks completed in 1 min) for the following problems.

a. Calculate the mean, standard deviation and 5-number summary for the following sets of data.

Hour 1

33, 51, 34, 48, 43, 53, 64, 55, 51, 47, 39, 34, 29, 38, 40, 65, 66, 60, 51, 65

29 Minimum
38.5 Q_1
49.5 Median
57.5 Q_3
66 Maximum
48.3 Mean
11.8 Std Deviation

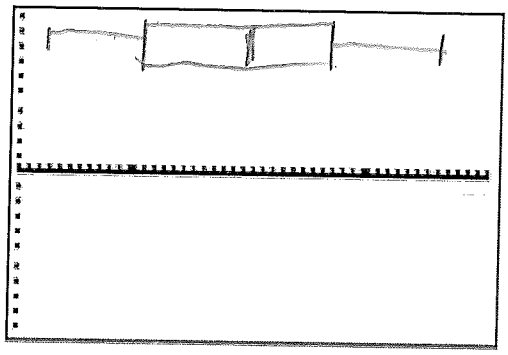
Hour 2

60, 55, 50, 51, 50, 48, 43, 53, 44, 51, 51, 47, 58, 34, 29, 19, 40, 12, 57, 50, 43

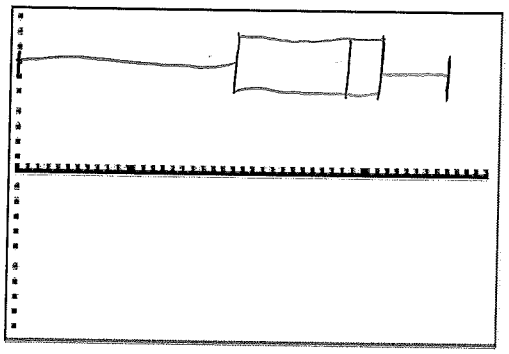
12 Minimum
41.5 Q_1
50 Median
52 Q_3
60 Maximum
45 Mean
12.4 Std Deviation

b. Use the graphing calculator to make a boxplot of the jumping jack data for Hour 1 and Hour 2 then sketch them here.

Hour 1



Hour 2



c. Which numerical measures *should* be used to describe the Hour 2 data: the *5-number summary* or the *mean and standard deviation*? Why?

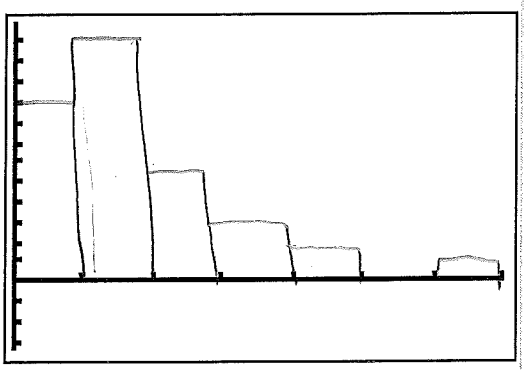
5-number summary because the data is skewed.

d. Which class was better at jumping jacks, Hour 1 or Hour 2? Write a sentence or two to justify your answer.

2. Use the data collected (Average # of pages read per week) to answer the following questions.

14, 5, 35, 50, 12, 28, 25, 50, 50, 75, 29, 33, 15, 17, 22, 77, 60, 55, 18, 14, 10, 0, 5, 19, 22, 27, 38, 51, 100, 21

a. Make a histogram of the data using your graphing calculator. Sketch what you see in the box provided.



b. Describe the distribution.

S: 0 to 100
O: 100
C: Med=26
S: Skewed right.

c. Which numerical measures *should* be used to describe the data: the *5-number summary* or the *mean and standard deviation*? Why?

5-# summary because the data is skewed.