

Warmup:

A student scored a 79, 86, 83, and a 91 on the first four tests of the semester. The student wants to increase her average to a 90. What would be the score necessary on the fifth test to do this?

$$79 + 86 + 83 + 91 = 339 + x$$

$$\frac{339 + x}{5} = 90 \cdot 5$$

$$339 + x = 450$$

$$-339 \quad -339$$

$$x = 111$$

$$90 \times 5 = 450$$

$$\begin{array}{r} 450 \\ -339 \\ \hline 111 \end{array}$$

If it were possible to score that high on the fifth test, what would be the mean, median, and mode of her five test scores?

$$\frac{90}{1}$$

90 79, 83, 86, 91, 111 No mode

1) Number of points scored by a basketball player during each game in a season:

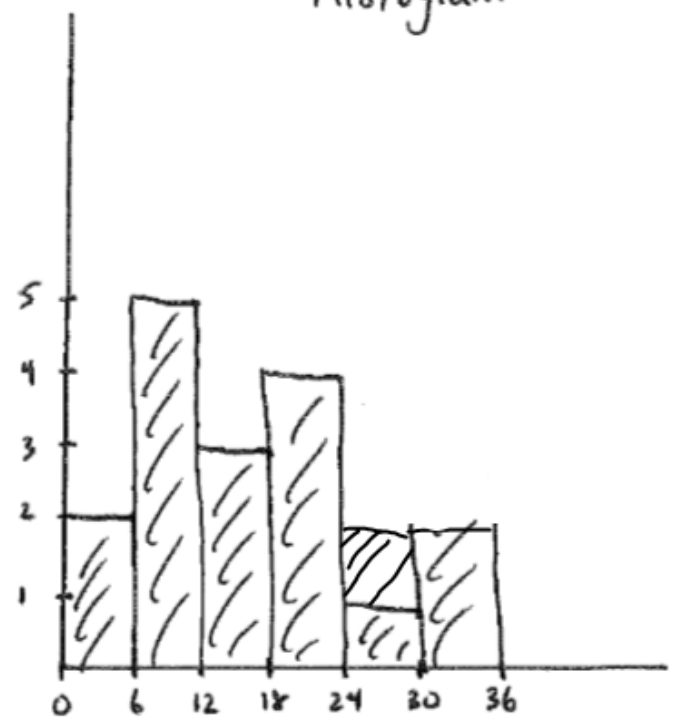
18 8 22 14 3 8 14 31 29
~~8~~ ~~33~~ ~~28~~ ~~19~~ ~~13~~ ~~6~~ ~~10~~ ~~20~~ ~~7~~

$$\frac{33-0}{6} = \frac{33}{6} = 5.5 = 76$$

Stem	Leaf
0	0, 3, 6, 7, 8, 9
1	0, 3, 4, 4, 8, 9
2	0, 2, 8, 9
3	1, 3

Freq. Dist.	
Pts. Scored	Freq.
0-5	2
6-11	5
12-17	3
18-23	4
24-29	2
30-35	2

Histogram



key: 1/3 = 13

2) Scores of 40 students on an exam:

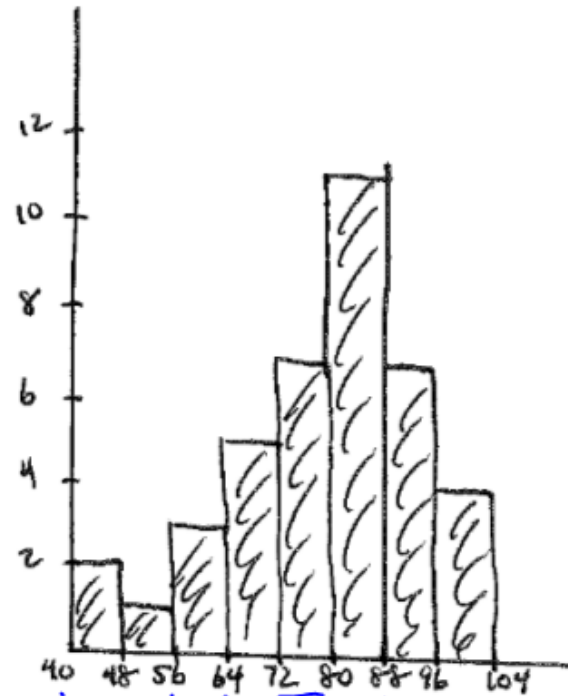
$$\frac{100 - 43}{8} = \frac{57}{8} = 7.125 \approx 7.8$$

85 100 90 82 76 64 55 70 86 93 79 80 62 43
 91 83 75 94 87 67 76 45 98 84 85 71 78 78
 88 68 57 100 97 81 73 71 84 87 92 95

Freq. Dist.

4 | 3, 5
 5 | 5, 7
 6 | 2, 3, 4, 7
 7 | 0, 1, 1, 3, 5, 6, 6, 8, 8, 9
 8 | 0, 1, 2, 3, 4, 4, 5, 5, 6, 7, 7, 8,
 9 | 0, 1, 2, 3, 4, 5, 7, 8
 10 | 0, 0 key 9/0 = 90

Exam Score	Freq.
40-47	2
48-55	1
56-63	3
64-71	5
72-79	7
80-87	11
88-95	7
96-103	4



UGA's final scores in each game.

B) Find the mean, median and mode of the opposing team's scores.

UGA Football 2010 season:

Opponent:	Result:
Louisiana-Lafayette	Won 55-7
South Carolina	Lost 17-6
Arkansas	Lost 31-24
Mississippi State	Lost 24-12
Colorado	Lost 29-27
Tennessee	Won 41-14
Vanderbilt	Won 43-0
Kentucky	Won 44-31
Florida	Lost 34-31
Idaho State	Won 55-7
Auburn	Lost 49-31
Georgia Tech	Won 42-34
UCF	Lost 10-6

UGA

55, 6, 24, 12, 27, 44, 48, 44, 31, 55, 31, 42, 6

in order: 6, 6, 12, 24, 27, 31, 31, 41, 42, 43, 44, 55, 55

Opponents

7, 17, 81, 24, 29, 14, 0, 31, 34, 7, 49, 34, 10

in order: 0, 7, 7, 10, 14, 17, 24, 29, 31, 31, 34, 34, 49

UGA mean

$$\frac{\text{sum}}{13} = 32.08$$

$$\frac{417}{13}$$

opponents mean

$$\frac{\text{sum}}{13} = \frac{287}{13} = 22.08$$

UGA

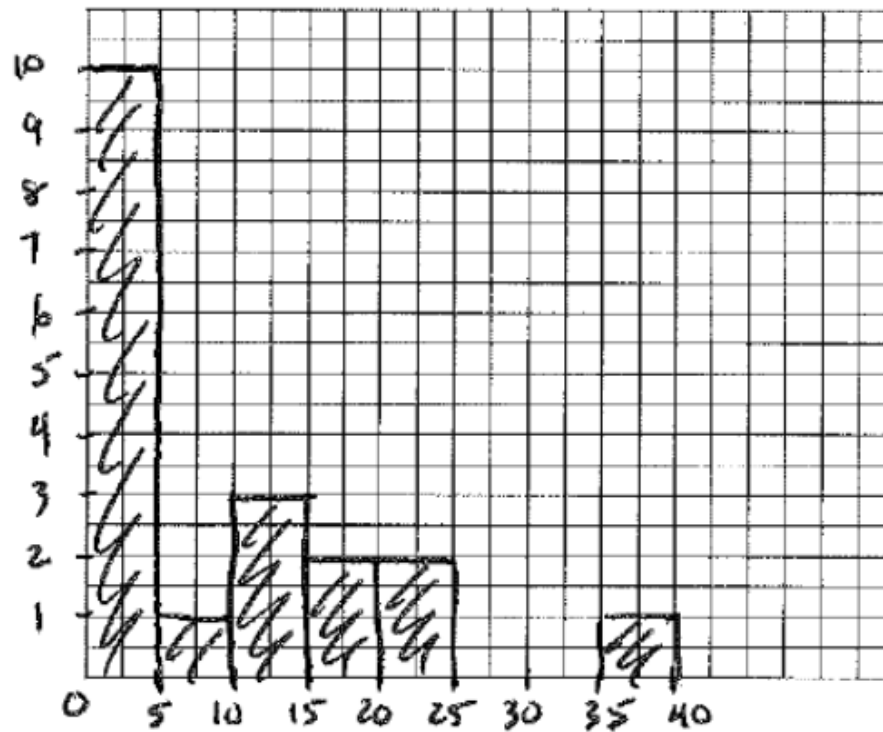
Mean: 32.08 Mode: 6, 31, 55 Median: 31

Opponents

Mean: 22.08 Mode: 7, 31, 34 Median: 24

4) Use the frequency distribution to draw a

Intervals for number of home runs	Number of players that scored home runs in the given interval
0-4 homeruns	TTTT TTTT
5-9 homeruns	
10-14	
15-19	
20-24	
25-29	
30-34	
35-39	↓



EQ: How can I use percentiles and a box and whisker plot to analyze data?

Percentiles:

The **Nth percentile** of a distribution is a value such that N% of the data falls at or below N and (100-N%) of the data fall at or above it.

You took an English achievement test to obtain college credit in freshman English by examination.

- 1) If your score was at the 89th percentile, what percentage of scores are at or below yours?

At or above yours?

11%

89%



- 2) If 4000 students took the achievement test, how many students scored lower than you?
How many students scored higher than you?

$$440 \leftarrow 4000 \times .11$$

$$\begin{array}{r} 4000 \\ - 3560 \\ \hline 440 \end{array}$$

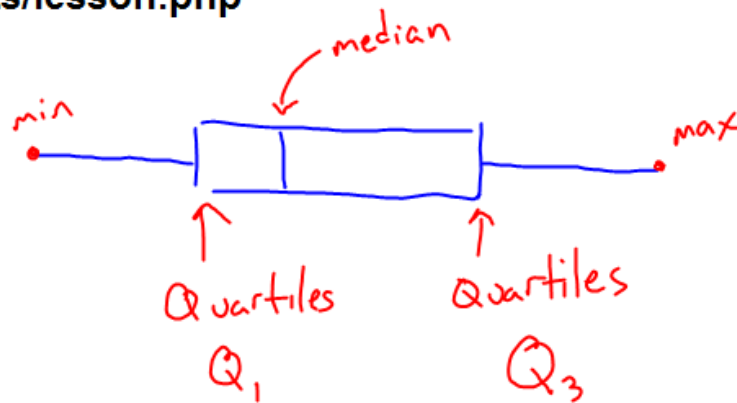
$$(4000)(.89) = 3560$$

- 3) If the scores ranged from 1 to 100 and your raw score is 95, does this mean that your score is at the 95th percentile?

No.

What is a Box & Whisker Plot?

<https://www.brainingcamp.com/content/box-and-whisker-plots/lesson.php>



Five Number Summary:

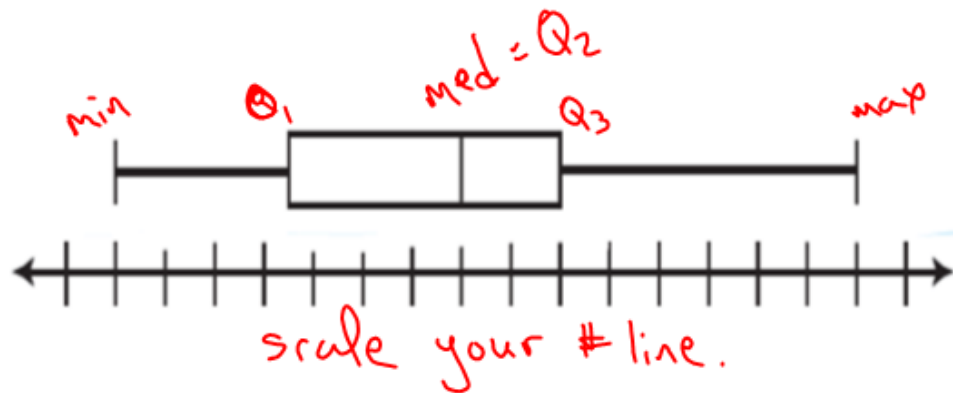
The five number summary consists of the **minimum data value**, **first (lower) quartile**, **median**, **third (upper) quartile**, and the **maximum data value**.

First Quartile: The median of the lower half of the data

Third Quartile: The median of the upper half of the data.

Box and Whisker Plot:

The five number summaries can be visually organized as a box and whisker plot.



The **middle 50%** of the data falls between the quartiles. The difference between the quartiles will give you an idea at how much the middle of your data is spread out. This difference is known as the **interquartile range**.

$$\rightarrow IQR = Q_3 - Q_1$$

What does a Box & Whisker plot tell us about the data?

A boxplot can show whether a data set is symmetric (roughly the same on each side when cut down the middle) or skewed (lopsided). A symmetric data set shows the median roughly in the middle of the box.

The median, part of the five-number summary, is shown by the line that cuts through the box in the boxplot.

Skewed data show a lopsided boxplot, where the median cuts the box into two unequal pieces. If the longer part of the box is to the right (or above) the median, the data is said to be *skewed right*. If the longer part is to the left (or below) the median, the data is *skewed left*.

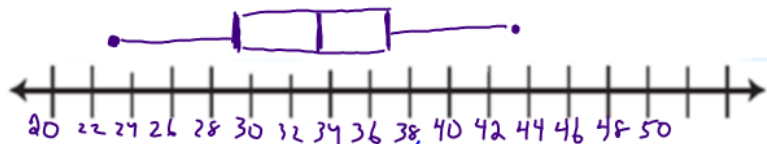
EQ: How can I use percentiles and a box and whisker plot to analyze data?

Example 1: The development office of a local college did a salary survey of alumni who graduated 2 years ago and have jobs. Sixteen alumni responded to the survey the first week. Their annual salaries (in thousands of dollars) are displayed below. Create a box and whisker plot for the data.

Annual Salaries:

38.5	39.5	32.0	30.5	36.8	29.2	23.7	34.1
28.3	27.9	33.6	37.0	43.5	34.6	33.8	36.1

Min: 23.7 Q1: 29.85 Median: 33.95 Q3: 36.9 Max: 43.5 IQR: 7.05
 $Q_3 - Q_1$



23.7, 27.9, 28.3, 29.85 } 30.5, 32, 33.6, 33.8 } 34.1, 34.6, 36.1, 36.8 } 37, 38.5, 39.5, 43.5

, 38.5, 39.5, 43.5

no. 1

Outliers:

Sometimes when working with data, not all of the data values fit the distribution. An **outlier** is a data value that is **significantly greater or lesser** than the other data values.

✦ -Outliers can often affect the statistics you calculate, **such as the mean.** ✦

-An outlier is found by calculating the **lower and upper fences**. Any value outside of these is considered an outlier.

✦ **Lower Fence:**

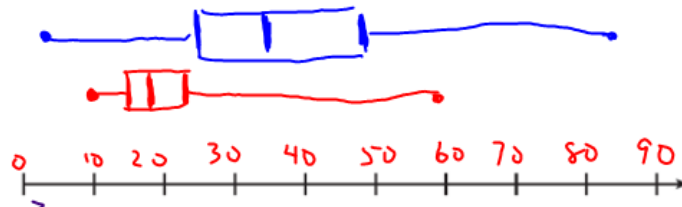
Upper Fence: ✦

$$Q1 - (IQR \cdot 1.5)$$

$$Q3 + (IQR \cdot 1.5)$$

Example 2: Brenda needs to get the oil changed in her car, but she hates to wait. Quick Change and Speedy Oil are two garages near Brenda's house. Brenda chooses twelve customers at random from each garage to look at wait times.

Quick Change		Speedy Oil	
10	60	22	15
12	24	20	18
16	23	22	15



Q.C.: 10, 12, 15, 15, 16, 18, 20, 22, 22, 22, 24, 60

S.O.: 5, 24, 26, 30, 30, 32, 40, 45, 45, 55, 60, 85

outlier

60 is an outlier

IQR: $22.5 - 15 = 7.5$
 $Q_3 - Q_1$

IQR: $(7.5)(1.5) = 11.25$
 $\times 1.5$

L.F. $Q_1 - 11.25$
 $15 - 11.25 = 3.75$

U.F. $Q_3 + 11.25$
 $22.5 + 11.25 = 33.75$

IQR: $50 - 28 = 22$
 $Q_3 - Q_1$

$(22)(1.5) = 33$

L.F.: $28 - 33 = -5$

[5, 83]

U.F.: $50 + 33 = 83$

With the outliers

Wait Times (minutes)

Quick Change

10 60 22 15

12 24 20 18

16 23 22 15

Speedy Oil

5 60 45 24

40 26 55 30

32 85 45 30



Without the outliers

Wait Times (minutes)

Quick Change

10 60 22 15

12 24 20 18

16 23 22 15

Speedy Oil

5 60 45 24

40 26 55 30

32 85 45 30



How do outliers affect the data?

Example 3: The table below shows the populations of the 10 largest cities in Georgia (in the thousands).

City	Population
Atlanta	420.0
Augusta	195.8
Columbus	189.9
Savannah	136.3
Athens	115.5
Sandy Springs	93.9
Macon	91.4
Roswell	88.3
Albany	77.4
Johns Creek	76.7

1) Calculate the five number summary and IQR for the data.

2) Calculate the lower and upper fences. Are there any outliers for the data.

3) Create a Box and Whisker plot for the data:



4) Calculate the mean with the outliers in the data set and without the outliers in the data set.

Mean with outliers:

Mean without the outliers:

How did the outliers affect the mean? _____

Which better represents the data? _____

Example 4: Coach Petersen's Middletown 9th grade football team is having a tough season. The team is struggling to win games. He is analyzing the data from the past two seasons.

Points Scored 2011	10	13	17	20	22	24	24	27	28	29	35
Points Scored 2012	0	7	17	17	18	24	24	24	25	27	45

a. Calculate the five number summary for each year.

b. Find the upper and lower fence for each set of data. Determine the outliers in each set of data.

c. Construct a box and whisker plot of each year's scores using the same number line.



d. Calculate the mean for each of the football seasons both with the outliers and without the outliers.

HW #2:

Box and Whisker
Plots