

Measures of Central Tendency

Mean – numerical average: $\text{mean} = \frac{\text{sum of the data values}}{\text{number of data values}}$

Median – middle value from a **sorted** list or mean of the middle two values

(the middle value from a list of n numbers is $\frac{n+1}{2}$)

Mode – the most frequently occurring value

(can have no mode, one mode, two modes (bimodal), three modes, (trimodal), etc.)

Example: Test scores: 78, 87, 84, 75, 80, 98, 78, 95, 72

$$\text{Mean} = \frac{78+87+84+75+80+98+78+95+72}{9} \approx 83$$

Median: first put the data in order:	72
	75
Find the middle number	78
	78
Median = 80	80
	84
	87
	95
	98

Mode: 78

Example: Heart Rates: count the number of times your heart beats in 15 seconds and multiply by 4 to get your heart rate per minute.

Stem Plot



Outliers are values that are significantly different (larger or smaller) than the rest of the data.

Do any of the heart rates from our class appear to be outliers? Which ones?

Quartiles – values that separate the data into four parts.

Test Scores

72	First find the median
75	(this is also known as the second quartile or Q_2)
78	
78	Now find the middle value of the lower and upper halves of data
80	(these are the lower quartile and upper quartile)
84	(they are also called Q_1 and Q_3)
87	
95	
98	

5-Number Summary

Low =

Q_1 =

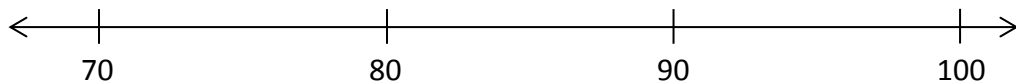
Median =

Q_3 =

High =

Box Plot (also known as a box-and-whisker plot)

- A plot that shows the information from the 5-number summary.
- It is useful for looking at the distribution of a set of values and also useful when comparing two or more sets of data



Construct a Box Plot for our class's Heart Rate data:

Range – the difference between the lowest value and the highest value

$$\text{Range} = \text{high} - \text{low}$$

Inter-Quartile Range (IQR) – the difference between the quartiles

$$\text{IQR} = Q_3 - Q_1$$

Modified Box Plot

- Uses a rule for finding outliers
- Defines an outlier as a value that is more than 1.5 x the IQR away from a quartile. (below the lower quartile or above the upper quartile)
- Outliers are shown on a modified box plot with a special symbol (like a *)
- The whiskers only extend to the lowest and highest values that are not outliers

Example – Test Score data

Test Scores

72	$Q_1 = 76.5$	
75	$Q_3 = 91$	
78		
78	Find the IQR	
80	$\text{IQR} = Q_3 - Q_1 = 91 - 76.5 = 14.5$	
84		
87	$1.5 \times \text{IQR} = 1.5 \times 14.5 = 21.75$	
95		
98	$Q_1 = 76.5$	$Q_3 = 91$
	<u>- 21.75</u>	<u>+ 21.75</u>
	54.75	112.75

Any score below 54.75 or above 112.75 is an outlier. We have no outliers in this data set.

Are there any outliers in the Heart Rate data set? Which values are they?

Construct a Modified Box Plot of the Heart Rate data set.

Percentile – a value that divides the range of data into two parts such that the part below the percentile contains a given percent of the data

21 The value 64 is at what percentile?

24 There are 8 values that are less than 64 and 20 values altogether.

31

45 $\frac{8}{20} = 0.40$ so 64 is at the 40th percentile

47

54

61

At what percentile is 98?

62

64

65

65

71

Find the value at the 25th percentile

82

$25\% \times 20$ values

87

$0.25 \times 20 = 5$

87

93

The value that has 5 numbers smaller than it is 54

97

54 is the value at the 25th percentile

98

98

98

Find the value at the 80th percentile

Example: Find the percentile for your heart rate using the Heart Rate data for our class.

Find the heart rate that is at the 50th percentile? What other name does this value have?