Name _____

Data with Outliers

R 3-2

An **outlier** is a number in a set of data that is very different from the rest of the data. Outliers can affect the measures of central tendency, particularly the mean.

Throughout most of the year, Acapulco, Mexico, is very sunny. If you look at the number of wet days from February through May, you see that the range is from 0 to 2. However, when the rainy season begins in June, the typical number of wet days per month jumps to 12. 12 is an outlier in this set of data because it is very different from the rest of the numbers.

Number of Wet Days in Acapulco, Mexico	
February	1
March	0
April	0
May	2
June	12

	Feb–May	Feb–June
Mean	$1 + 0 + 0 + 2 = 3$ $3 \div 4 = \underline{0.75}$	$1 + 0 + 0 + 2 + 12 = 15$ $15 \div 5 = \underline{3}$
Median	0, $\underbrace{0, 1,}_{\underline{0.5}}$ 2	0, 0, $\underline{1}$, 2, 12
Mode	$\underline{0}$, $\underline{0}$, 1, 2	$\underline{0, 0}$, 1, 2, 12

The outlier did not affect the mode, and it changed the median slightly. But look what happens to the mean when the June number is included in the data. The mean becomes 3 wet days, which is the same as February through May combined!

Identify the outlier. Then find the mean, median, and mode of the data with and without the outlier.

1. 190, 210, 160, 250, 1,400, 190

Which number is the outlier? _____

	With Outlier	Without Outlier
Mean		
Median		
Mode		

2. 86, 77, 20, 81, 70, 86

Which number is the outlier? _____

	With Outlier	Without Outlier
Mean		
Median		
Mode		