Introduction to Statistics
Activity 2.3b **SOLUTIONS**

State Populations

a) The population of the United States grew by 9.1% between 2000 and 2009. The following table reports the percentage growth in population for each of the 50 states during this decade. States are classified by whether they are east or west of the Mississippi River:

Western State	%	Western State	%	Eastern State	%	Eastern State	%
Alaska	11.4	Montana	8.1	Alabama	5.9	New Hampshire	7.2
Arizona	28.6	Nebraska	5.0	Connecticut	3.3	New Jersey	3.5
Arkansas	8.1	Nevada	32.3	Delaware	13.0	New York	3.0
California	9.1	New Mexico	10.5	Florida	16.0	North Carolina	16.6
Colorado	16.8	North Dakota	0.7	Georgia	20.1	Ohio	1.7
Hawaii	6.9	Oklahoma	6.9	Illinois	4.0	Pennsylvania	2.6
Idaho	19.5	Oregon	11.8	Indiana	5.6	Rhode Island	0.5
Iowa	2.8	South Dakota	7.6	Kentucky	6.7	South Carolina	13.7
Kansas	4.8	Texas	18.8	Maine	3.4	Tennessee	10.7
Louisiana	0.5	Utah	24.7	Maryland	7.6	Vermont	2.1
Minnesota	7.0	Washington	13.1	Massachusetts	3.9	Virginia	11.4
Missouri	7.0	Wyoming	10.2	Michigan	0.3	West Virginia	0.6
				Mississippi	3.8	Wisconsin	5.4

b) Consider the western states. Would it be reasonable to use 0.1 as the leaf unit? Explain. [*Hint*: How many stems would there be? How many leaves would be on each stem?]
No, this would produce 33 stems (0,1,..,32) for 24 leaves. Ideally we want a number of stems such that there are more leaves than stems. Visually, this will help reveal any patterns in the data. Too many stems means that few stems have more than one leaf.

c) Still considering the western states, would it be reasonable to use 1 as the leaf unit? Explain.
Certainly, for this we would use 10s place as the stems. This would provide us four stems, with many leaves appearing on each. Seems like too few.

In creating stemplots, it is most effective to have approximately 5–15 stems. Having too many stems may make it difficult to see the overall pattern (e.g., clustering). Too few stems and important details can be lost because the data are clumped together without showing much pattern (e.g., gaps, skewness). One way to change the level of detail displayed is to create a **split stemplot** where each stem appears twice, the 0–4 leaves appear on the first stem and the 5–9 leaves appear on the second. When you want fewer stems, you can ignore the last digit in the data as we did in part c.

d) If we create a split stemplot for the western states, how many stems will we have?
Doubling the stems from part B, we would get 8 stems.

To investigate whether eastern or western states tend to have higher population growths, you want a display that helps directly compare the two groups. In a **side-by-side stemplot**, a common set of stems (for example, the ones digit) is placed in the middle of the display with leaves (for example, the tenths digit) branching out in either direction to the left and right. The convention is to order the leaves from the middle out from least to greatest.

e) Complete the following side-by-side stemplot to compare the distributions of population growth between eastern and western states by filling in the states beginning with the letter *N*. [*Hints*: Use the stems provided here. An underscore (_) indicates where you should fill in a leaf value. Use single digit stems by truncating the tenths digit after the decimal point. So, Alabama's percentage appears as 1|1 ignoring the 0.4.]

Western states		Eastern states
42_0	0	000122_ _33334
98877766_	0	5556_7
311_0	1	0133
986	1	_6
4	2	0
8	2	
_	3	
	3	

Leaf unit = one percentage point

f) Based on this side-by-side stemplot, compare the distributions of population growth between eastern and western states. [*Hint*: Remember to comment on center, spread, shape, and any other apparent features, such as outliers, identifying the unusual states by name.]
Both graphs are skew right, with most of their data less than 10%. However, we see the maximum for the Western states is a bit higher than for the Eastern states. We also see that the data for the Western states is spread out over 7 stems, as opposed to only 5 for the Eastern states. There are no empty stems between any of the data points, so there aren't really any obvious outliers. Although Nevada at 32.3 is pretty big.