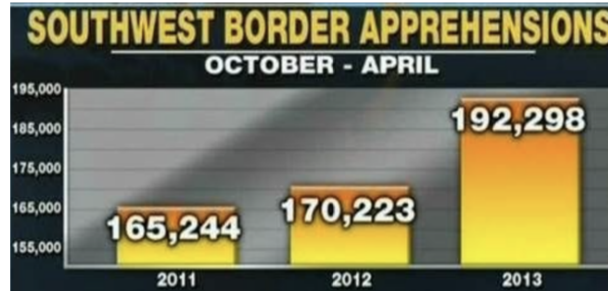**Deep Thoughts**
Unit 1: One Variable Data

1. The following graph was displayed by a national news organization. Explain why the graph may be misleading, and sketch a corrected version of the graph.

2. A real estate agent is collecting data on the number of houses built in his towns three neighbourhoods during three different decades. The table below gives information.

|	1960s	1970s	1980s
Shady Lane	40	30	10
Oakcrest	60	15	5
Pinewood Estates	0	45	15

(a) What proportion of houses were built in Pinewood Estates?

(b) Find the distribution of Decade Built for the houses in this town using relative frequencies.

(c) What percent of homes were built in Oakcrest and in the 1960s?

3. The following table gives the result of a random sample of upper level students at Rocky Vista University (the Fighting Prarie Dogs!), along with a mosaic plot.

| Employment Status	Grade Level	
	Junior	Senior
Currently working	14	30
Not working but have had a job	22	40
Never had a job	15	10

Employment status
- Never had a job
- Not working but did in the past
- Currently working
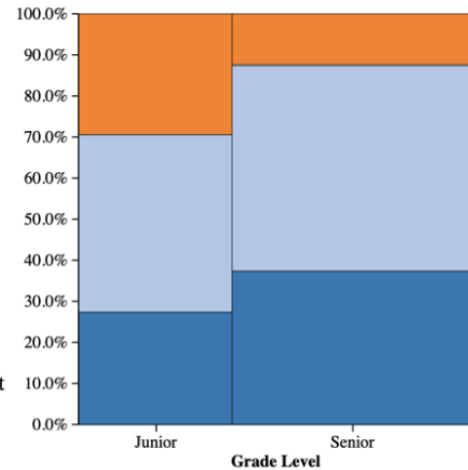
(a) Calculate the proportion of Junior that are currently working, not working but have had a job, and never had a job.

(b) Calculate the proportion of Seniors that are currently working, not working but have had a job, and never had a job.

(c) Write a few sentences summarizing what the display in part (a) reveals about the association between grade level and job experience for the students the sample.

4. Mr. Kevin is a huge fan of University of Michigan football. His favourite season was the 1997 season. The dotplot shows the number of points scored by the U of M team in the 12 games that season.

Points Scored
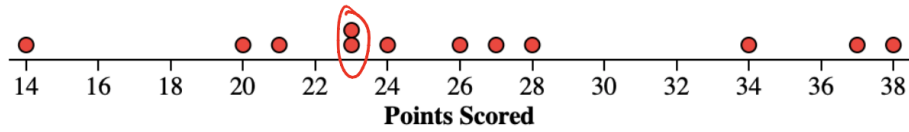
(a) Use the dots to create a stemplot of the distribution.

(b) Are there any potential outliers? Why?

(c) What measure of center is the most appropriate to describe the distribution? Explain.

5. A researcher is interested in how much annual rainfall is typical in the United States. She takes a random sample of 9 cities in the U.s. and records the annual rainfall, in inches.

8.2
10.3
33.5
39.1
40.5
41.9
42.4
44.9
53.7
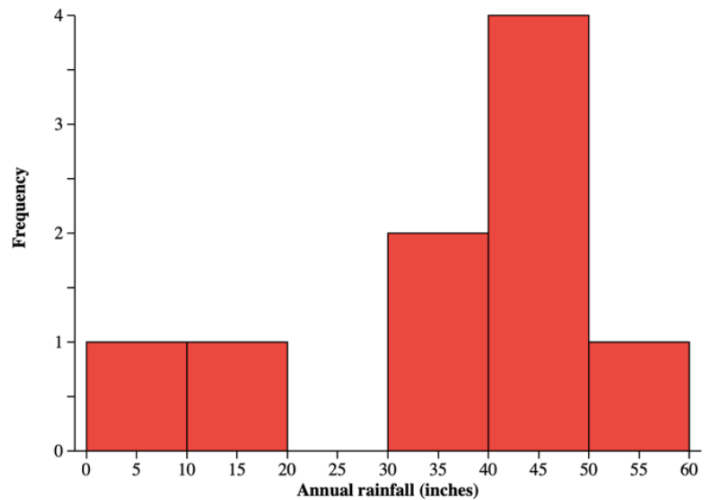
(a) Calculate the mean annual rainfall for these cities.

(b) Find the median annual rainfall for these cities.

(c) Would you use the mean or the median to summarize the typical annual rainfall for a U.S. city? Explain

(d) The standard deviation of the annual rainfall for these 9 cities is 15.52 inches. Interpret this value.

6. Mr. Wilcox is a huge fan of University of Michigan football. His favorite season was the 1997 season (a perfect season!). Here is a back-to-back boxplot of the points scored by the 1997 University of Michigan football team and the archrival Michigan State University football team. Write a few sentences comparing the distributions.
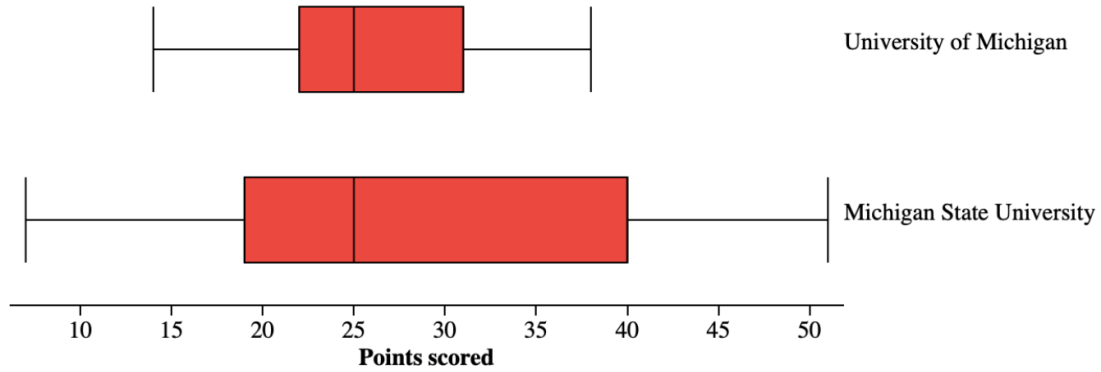
Deep Thoughts
Unit 1: One Variable Data

1. According to a 2019 article at inder.com. the state of Pennsylvania was at the 82nd percentile for Pre-K to 12th grade education and was at the 0th percentile for higher education. Explain what these values mean.

2. The graph displays the cumulative relative frequency of the cost of in-state public college education for each of the 50 states.
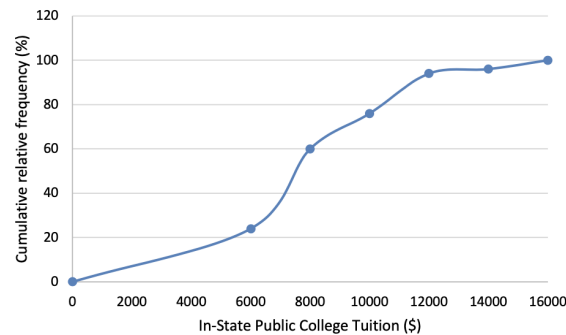
   (a) About what percent of states have in state public college tuition less than or equal to $8000? More than $8000?

   (b) Estimate Q_1, Q_3, and the IQR of the distribution of phone in state public tuition.

3. According to an article at Yahoo news, you should change your sheets every 7 days at minimum. To investigate the sheet changing habits of adults, a random sample of 20 adults reported how often they change their sheets using an anonymous survey. Here is a dotplot and summary statistics of the results.

Time before changing sheets (days)
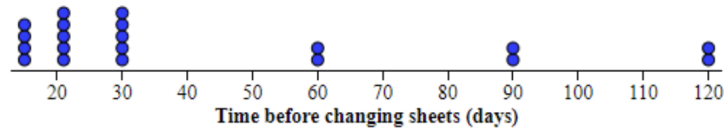
n	mean	SD	min	Q_1	med	Q_3	max
20	42.75	34.984	15	21	30	60	120

(a) Suppose you convert the time before changing sheets from days to weeks.Describe the shape, mean, and standard deviation of the distribution of time before changing sheets in weeks.

(b) The adults in the study are given an article explaining the health benefits that would arise from changing their sheets more often. After reading the article each person agrees to change their sheets one week sooner than they used to. How does the shape, center, and variability, of this distribution compare with the distribution of time in part (a)?

(c) Now suppose you convert the time before changing the sheets from part (b) to z scores. What would be the shape, mean, and standard deviation of the distribution. Explain your answers.