

Module 1 Challenge Problems

1.	Determine what the key terms (refer to in the following study. We want to know the average (mean) amount of money spent on school uniforms each year by families with children at Knoll Academy. We randomly survey 100 families with children in the school. Three of the families spent \$65, \$75, and \$95, respectively. KEY TERMS: Population _____ Statistic _____ Parameter _____ Sample _____ Variable _____ Data _____
2.	Consider the patients who went to the Urgent Care at the local hospital last week as the observational units (people we are studying) in a statistical study. a) Identify two categorical/qualitative variables that could be recorded on these observational units. b) Identify two numerical/quantitative variables that could be recorded on these observational units. c) State a research question that you could investigate about these observational units, using at least one of the variables that you gave in part a) and b).
3.	Imagine you work for a large and busy hotel, and you would like to perform a survey on the guests staying at the hotel throughout the year. Now, a census may work, but let's say we have limited resources so we would like to use a random sampling method to perform the survey. a) How might you sample the guests using a stratified random sample method? Be sure to identify what variable you would use as your strata and why. Describe the specific step-by-step process you would use to randomly select the guests for your survey. b) How might you sample the guests using a cluster sample method? Be sure to identify what variable you would use as your cluster and why. Describe the specific step-by-step process you would use to randomly select the guests for your survey.
4.	Considering yourselves to be observational units, explain why the following are not variables: • Average amount of hours worked in the past 24 hours among students in our class. • Proportion of students in our class who are first-generation college students • Have left-handers read more Harry Potter books, on average, compared to right-handers?

5.

The table shows the motor vehicle deaths in the United States by year from 2010 through 2018.

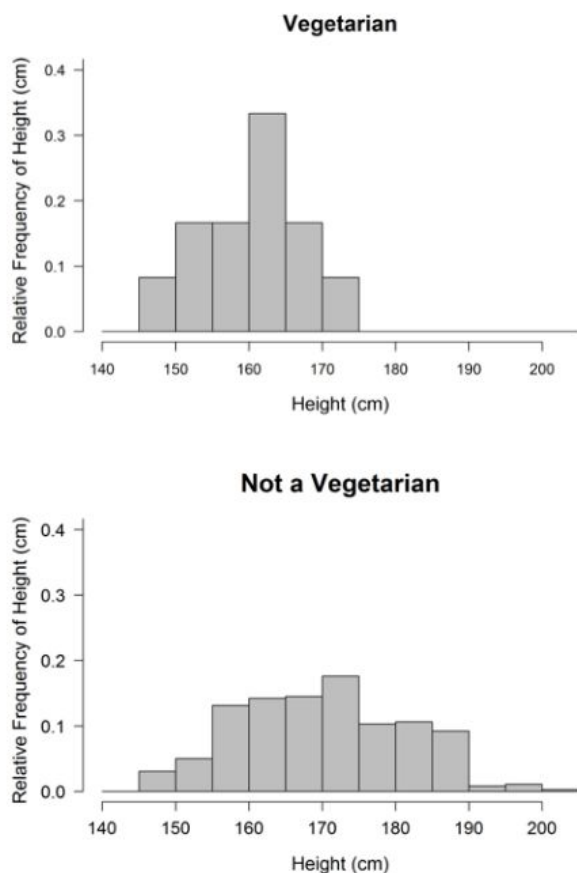
- What is the frequency of deaths measured from 2010 through 2018?
- What percentage of deaths occurred after 2016?
- What is the relative frequency of deaths that occurred in 2012 or before?
- What is the percentage of deaths that occurred in 2015?
- What is the cumulative relative frequency for 2016? Explain what this number tells you about the data.

Year	Deaths
2010	32899
2011	32479
2012	33782
2013	32893
2014	32744
2015	35425
2016	37806
2017	37133
2018	36560

Show your work below. Round relative and cumulative frequencies to 3 decimal places and percents to 1 decimal place.

Module 2 Challenge Problems

1.



A random sample of 411 high school students was selected in order to compare the heights, in centimeters (cm), of vegetarians and nonvegetarians. The relative-frequency histograms below show the distributions of height for the 15 students who said that they were vegetarian and for 396 students who said they were not vegetarian.

- What are the observational units? What are the variables, and what type of variable are they?
- Write a few sentences comparing the distributions of height for the vegetarians and nonvegetarians.
- Explain why it is better to use relative frequencies (proportions) rather than frequencies (counts) when comparing the vegetarians and nonvegetarians.
- Based on the design of this would it be reasonable to conclude that being a vegetarian causes students to be shorter, on average? Explain.

2.

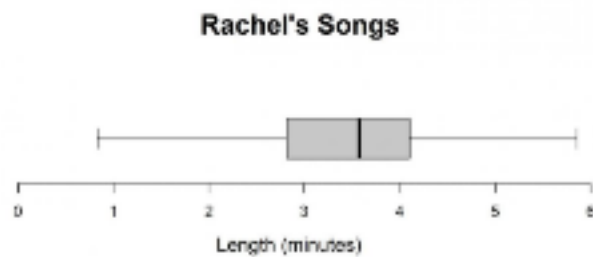
Use the following information to answer the next six exercises. Sixty-five randomly selected car salespersons were asked the number of cars they generally sell in one week. 14 people answered that they generally sell three cars; 19 generally sell four cars; 12 generally sell five cars; 9 generally sell six cars; 11 generally sell seven cars.

First quartile = _____, Second quartile = median = 50th percentile = _____

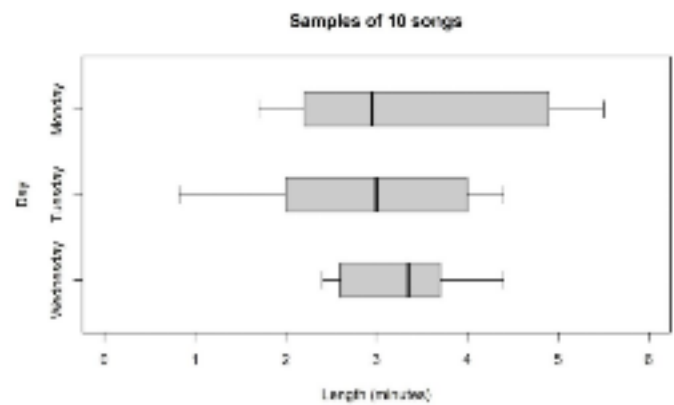
Third quartile = _____, Interquartile range (*IQR*) = _____ – _____ = _____

10th percentile = _____, 70th percentile = _____

3. Rachel has a large number of songs on her music player. The song lengths, in minutes, are summarized in the boxplot below.



Rachel decided to listen to ten songs each morning. She set the music player to select the songs at random. The boxplots below summarize song lengths for Monday, Tuesday, and Wednesday of that week.



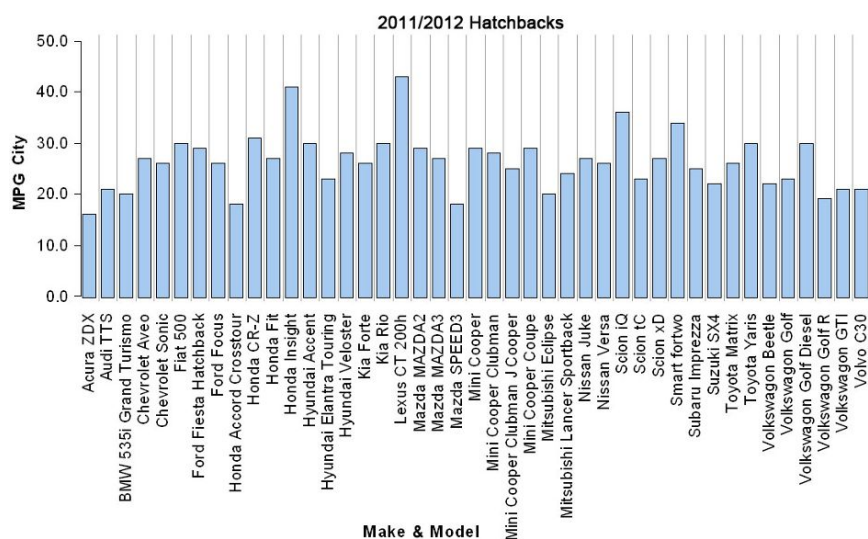
Of the following 4 statements - 2 of them are true. Identify which are true. Then write 2 sentences explaining your choices.

True or False	The song lengths varied less on Wednesday than the song lengths on Monday or Tuesday.	True or False	Something must be wrong with the music player, because the boxplots for Monday, Tuesday, and Wednesday all look very different.-
True or False	At least 55% of the songs will have a song length greater than 3 minutes.	True or False	The song lengths for Monday will have a mean less than median.

Module 3 Challenge Problems

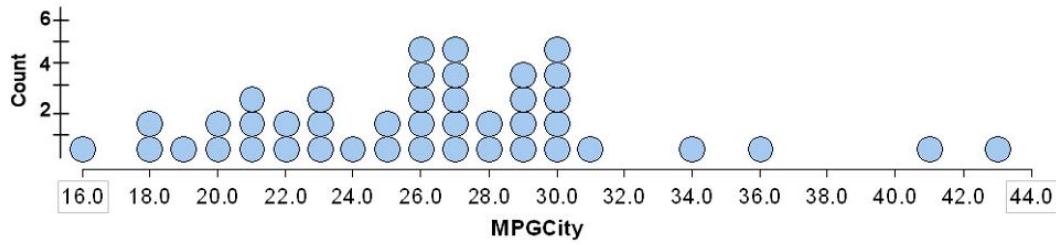
- 1.
- a) For each of the following quantities, indicate whether it can NEVER be negative or can SOMETIMES be negative.
- Standard deviation
 - Interquartile range
 - z-score
- b) Which would be larger – the standard deviation of the weights of 1000 randomly selected people, or the standard deviation of the weights of 10 randomly selected cats (ordinary domestic cats)?

2.



- a) What type of graph is this?
- b) What variables are represented in the graph? What types are they?
- c) What is/are the mpg rate(s) that occur most frequently (this measure of center is the mode)?

3.



- How many hatchbacks get 23 mpg in the city? Describe the distribution of the dotplot.
- What is/are the mpg rate(s) that occur most frequently?
- Which mpg rate is the middle rate (this measure of center is typically called the median)? What is the mean mpg?
- Take the data from the dotplot to *create a boxplot, stem-and-leaf plot, and histogram*. You must use software, such as StatCrunch or the Art of Stat to create your graphs.

4.

Considering the MPGCity data set for hatchback cars, would you be surprised if a company made a hatchback that could get 55 mpg? [Hint: Consider if that value is an outlier]