

FST NOTES 1-7

TOPIC: Comparing Numerical Distributions

GOAL

Write descriptions of the differences between two data sets as seen from their distributions.

SPUR Objective

G Use statistics to compare and contrast two data sets in the context of a situation.

VOCABULARY: NOTHING NEW.

The previous lessons used statistics to describe distributions of numerical variables and to develop cumulative distributions. This lesson asks you to compare two or more distributions. The emphasis is on clear, descriptive writing.

A written description of distributions compares and contrasts their *shapes, centers, and spreads*. The description should use the vocabulary and statistical measures you have learned in prior lessons. Although “a picture is worth a thousand words,” you should not need that many words. Your description should be clear enough so that a reader who does not have graphs of the distributions can create mental images of them.

NOTES

When comparing statistical summaries, notice how many times larger one number is than another. Make note of where the greatest relative differences are. Besides looking at the values of the 5-number summary in the box plots, examine where a quartile in one distribution might be in the other distribution. Discuss outliers, if possible. The histograms give a better picture of the shape of the distribution. Information from several types of displays of two distributions can help you to give a more complete comparison of them.

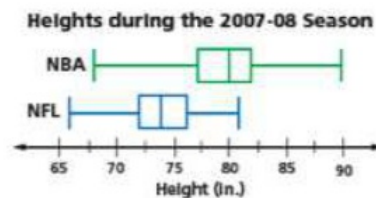
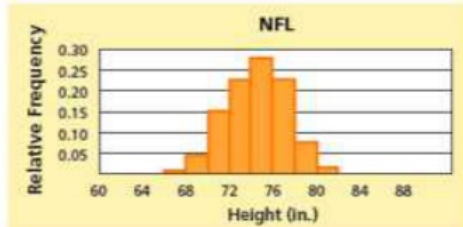
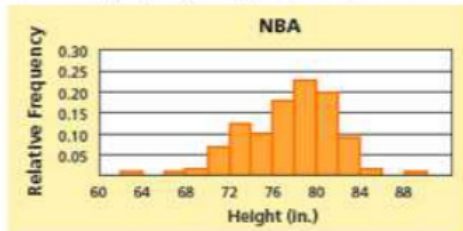
As you practice writing this kind of statistical summary, here is a checklist of things to remember.

- Describe the variable involved.
- Address all three major topics: shape, center, and spread.
- Give the appropriate units.
- Mention both data sets in your comparisons.
- Use words to give descriptions of your numbers. Use numbers to make your words more specific.

Activity 1

Below and at the right are histograms, statistical summaries, and box plots of the distribution of heights (in inches) of all players in the National Basketball Association (NBA) and the National Football League (NFL) during the 2007-2008 season.

Statistics	Heights (in inches)	
	NBA	NFL
mean	79.2	74.0
median	80	74
interquartile range	5	4
standard deviation	3.63	2.58
range	25	15



- Step 1 Write a sentence about one piece of information conveyed by the statistical summaries table to compare NBA and NFL player heights.
- Step 2 Write one sentence describing a comparison of NBA and NFL player heights that the box plots tell you that is not found in the table.
- Step 3 Write one sentence comparing NFL and NBA player heights using a piece of information from the two histograms that is not found in either the box plots or the statistical table.

Shape (skewed, symmetric, tail)

NFL - heights are symmetric

NBA - there are tails on both ends

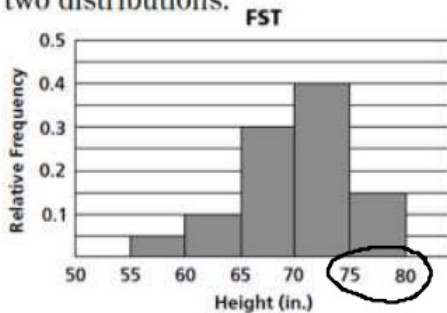
Center (mean, median)

NBA players on average are taller than NFL players, 5 inches taller

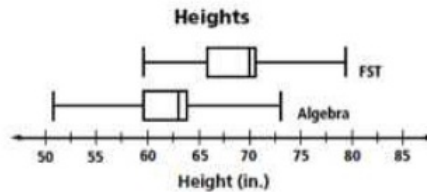
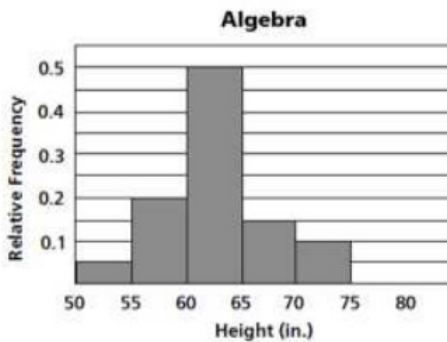
Spread (IQR, Range, outlier, standard deviation, variance)
 Range of heights much greater in the NBA
 S.D. larger for NBA - greater variability

Additional Example

The tables, histograms, and box plots below contain information about the heights (in inches) of 20 students in an FST class and 20 students in an Algebra class. Write a paragraph comparing and contrasting the two distributions.



Statistics	Height (in.)	
	FST	Algebra
mean	69.3	62.65
median	70	63
interquartile range	5	5.5
standard deviation	4.73	5.09
range	19	21



Shape (skewed, symmetric, tail)

FST - is slightly skewed right

Algebra - is fairly symmetric

center (mean, median)

Algebra students are on average shorter than FST students, 7 inches shorter

spread (IQR, range, outlier, S.D., variance)

IQR, S.D., and the range are similar