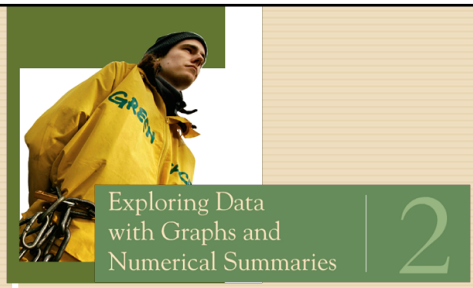


Exploring Data with Graphs and Numerical Summaries

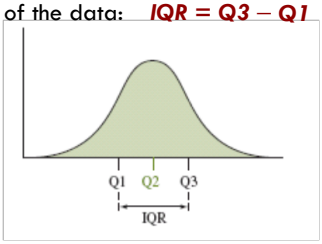
2

2.5 How Measures of Position Describe Spread

Aug 8-9:19 AM

Calculating Interquartile Range

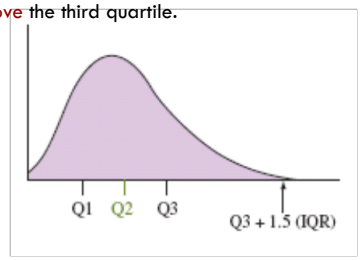
The interquartile range is the distance between the thirdand first quartile, giving spread of middle 50% of the data: ***IQR = Q3 – Q1***



Aug 8-9:19 AM

Criteria for Identifying an Outlier

An observation is a potential outlier if it falls more than *1.5 x IQR* below the first or more than *1.5 x IQR* above the third quartile.

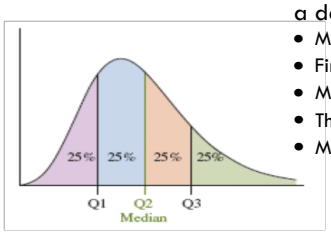


Aug 8-9:19 AM

5 Number Summary

The five-number summary of a dataset consists of:

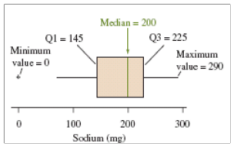
- Minimum value
- First Quartile
- Median
- Third Quartile
- Maximum value



Aug 8-9:19 AM

Boxplot

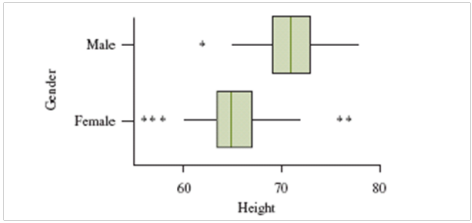
- Box goes from the Q1 to Q3
- Line is drawn inside the box at the median
- Line goes from lower end of box to smallest observation not a potential outlier and from upper end of box to largest observation not a potential outlier
- Potential outliers are shown separately, often with * or +



Aug 8-9:19 AM

Comparing Distributions

Boxplots do not display the shape of the distribution as clearly as histograms, but are useful for making graphical comparisons of two or more distributions



Aug 8-9:19 AM

Z-Score

z-score

The **z-score** for an observation is the number of standard deviations that it falls from the mean. For sample data, the z-score is calculated as

$$z = \frac{\text{observation} - \text{mean}}{\text{standard deviation}}.$$

An observation from a bell-shaped distribution is a potential outlier if its z-score < -3 or > +3

Aug 8-9:19 AM

2.79 **Female heights:** For the 261 female heights shown in the box plot in Figure 2.16, the mean was 65.3 inches and the standard deviation was 3.0 inches. The shortest person in this sample had a height of 56 inches.

a. Find the z-score for the height of 56 inches.

b. What does the negative sign for the z-score represent?

c. Is this observation a potential outlier according to the three standard deviation distance criterion? Explain.

Aug 8-9:19 AM

Aug 14-3:16 PM