WHAT HAVE WE LEARNED?

- We've learned the value of comparing groups and looking for patterns among groups and over time.
- We've seen that boxplots are very effective for comparing groups graphically. When we compare groups, we discuss their shape, center, and spreads, and any unusual features.
- We've experienced the value of identifying and investigating outliers. And we've seen that when we group data in different ways, it can allow different cases to emerge as possible outliers.
- We've graphed data that have been measured over time against a time axis and looked for long-term trends.

Terms

Boxplot

81. A boxplot displays the 5-number summary as a central box with whiskers that extend to the non-outlying data values. Boxplots are particularly effective for comparing groups and for displaying outliers.

Outlier

81, 87. Any point more than 1.5 IQR from either end of the box in a boxplot is nominated as an outlier.

Far Outlier

81. If a point is more than 3.0 IQR from either end of the box in a boxplot, it is nominated as a *far outlier*.

Comparing distributions

82. When comparing the distributions of several groups using histograms or stem-and-leaf displays, consider their:

- Shape
- Center
- Spread

Comparing boxplots

83. When comparing groups with boxplots:

- Compare the shapes. Do the boxes look symmetric or skewed? Are there differences between groups?
- Compare the medians. Which group has the higher center? Is there any pattern to the medians?
- Compare the IQRs. Which group is more spread out? Is there any pattern to how the IQRs change?
- Using the IQRs as a background measure of variation, do the medians seem to be different, or do they just vary much as you'd expect from the overall variation?
- Check for possible outliers. Identify them if you can and discuss why they might be unusual. Of course, correct them if you find that they are errors.

Timeplot

88. A timeplot displays data that change over time. Often, successive values are connected with lines to show trends more clearly. Sometimes a smooth curve is added to the plot to help show long-term patterns and trends.

Skills

- Be able to select a suitable display for comparing groups. Understand that histograms show distributions well, but are difficult to use when comparing more than two or three groups. Boxplots are more effective for comparing several groups, in part because they show much less information about the distribution of each group.
- Understand that how you group data can affect what kinds of patterns and relationships you are likely to see. Know how to select groupings to show the information that is important for your analysis.
- Be aware of the effects of skewness and outliers on measures of center and spread. Know how to select appropriate measures for comparing groups based on their displayed distributions.
- Understand that outliers can emerge at different groupings of data and that, whatever their source, they deserve special attention.
- Recognize when it is appropriate to make a timeplot.