M1.1 One Variable Statistics

• Create dot plots, histograms, and box plots.
• Use available classroom technology to create histograms and box plots and calculate measures of center and spread.
• Use terms such as "flat," "skewed," "bell-shaped," and "symmetric" to describe data distributions.
• Analyze and compare data sets.
• Understand relationships between mean and median for symmetrical and skewed data distributions.
• Recognize outliers when they exist, and know to investigate their source.
• Know that outliers affect the mean, but not the median of a data set.
• Describe variability by calculating deviations from the mean.
• Compare two data sets with the same means but different variabilities, and contrast them by calculating the deviation of each data point from the mean.
• Understand that IQR is a description of variability better suited to a skewed distribution.
• Work with two-way tables.

Instead of creating representations of data, the emphasis in high school is on interpreting representations and judiciously interpreting measures of center and spread. Students describe the shape of a data distribution in more detail (symmetric, skewed, flat, or bell-shaped). Students develop a more precise understanding of measures of center and understand relationships between mean and median for symmetrical and skewed data distributions. They learn that outliers affect the mean of a data set but not the median. They recognize outliers when they exist and learn to investigate their source. Students learn that standard deviation is a measure of spread, that a larger standard deviation means the data are more spread out, and to understand standard deviation as "typical distance from the mean" for a symmetrical distribution. They also

understand that interquartile range is a description of variability better-suited to a skewed distribution. Finally, students are introduced to two-way frequency tables and understand how to interpret relative frequencies in the context of the data represented in the tables.

In grades 6 to 8, students were introduced to data sets and different ways to represent data (histograms, dot plots, box plots). Statistics is introduced as a tool to answer questions about a population that have variability in the answer. Students learn about measures of center (median, mean) and measures of variability (interquartile range, mean absolute deviation), using them to draw informal comparative inferences about two populations. Students build on and expand their understandings of statistics in this unit. The key characteristics (measures of shape, center, and spread) are again seen and in addition, students may further describe the shape of a data distribution (symmetric, skewed, flat or bell shaped) and summarize by a statistic measuring center and a statistic measuring spread. Instead of creating representations of data, the emphasis in high school is on interpreting representations and judiciously interpreting measures of center and spread. Students develop more precise understanding of measures of center. They learn that mean and median are equal for symmetrical distributions, explain why mean and median are not equal in examples of skewed distributions, select median as the better measure of center for skewed distributions, and make generalizations about what kinds of distributions have means larger than medians and which have medians larger than means. Students learn that standard deviation is a measure of spread, that a larger standard deviation means the data are more variable or spread out, and the meaning of standard deviation as “typical distance from the mean” for a symmetrical distribution. To aid in developing their understanding, students will calculate a standard deviation by hand for a small data set at least once. Given different visual representations of data (box plots, histograms, dot plots) students draw and justify significant and meaningful conclusions about the given situation. (All of these representations are frequency graphs, however, the Standards do not require students to know or use the term “frequency graph,” although they use the term “frequency tables.”) Students are introduced to two-way frequency tables and understand how to interpret relative frequencies in the context of the data represented in the tables. In unit S2 (which could take place either before or after this unit), students also build their statistics foundation by learning ways to

determine whether two sets of data are correlated, and how strongly. Students identify linear association and interpret slope and intercept in the context of the data. Given different visual representations of data (linear models) students draw and justify significant and meaningful conclusions about the given situation. Students begin to use technology as a means to plot data and generate correlation coefficients. In S3, students revisit two-way frequency tables from a probability standpoint, and use them as a tool for conceptualizing and finding conditional probabilities. In S4, students combine the ideas of distributions and probability. They learn about normal distributions and use them to solve problems, and use the distributions of probability models to find the likelihood of a particular outcome. In doing so, students build on their experience with standard deviations from S1, calculating standard deviations using technology, and interpreting the results. Every high school statistics and probability standard is a modeling standard, hence modeling pervades the four units.

M1.1.0 Pre-unit diagnostic assessment

Assess students' recall of middle grades statistics, specifically their ability to

• recognize a statistical question;
• describe the distribution of data collected to answer a statistical question by its center, spread, and overall shape;
• interpret statistical plots;
• summarize numerical data sets in relation to their context;
• informally assess the degree of visual overlap of two numerical data distributions with similar variabilities.

Instead of creating representations of data, the emphasis in high school is on interpreting representations and judiciously interpreting measures of center and spread. Students describe the shape of a data distribution in more detail (symmetric, skewed, flat, or bell-shaped). Students develop a more precise understanding of measures of center and understand relationships between mean and median for symmetrical and skewed data distributions. They learn that outliers affect the mean of a data set but not the median. They recognize outliers when they exist and learn to

investigate their source. Students learn that standard deviation is a measure of spread, that a larger standard deviation means the data are more spread out, and to understand standard deviation as "typical distance from the mean" for a symmetrical distribution. They also understand that interquartile range is a description of variability better-suited to a skewed distribution. Finally, students are introduced to two-way frequency tables and understand how to interpret relative frequencies in the context of the data represented in the tables.

M1.1.1 How can data be represented and summarized meaningfully?

- **Revisit various ways to plot data: dot plots, histograms, and box plots.**
- **Interpret plots of data within the context of the data.**
- **Use the terms "symmetric" and "skewed" as descriptors of distributions.**

Instead of creating representations of data, the emphasis in high school is on interpreting representations and judiciously interpreting measures of center and spread. Students describe the shape of a data distribution in more detail (symmetric, skewed, flat, or bell-shaped). Students develop a more precise understanding of measures of center and understand relationships between mean and median for symmetrical and skewed data distributions. They learn that outliers affect the mean of a data set but not the median. They recognize outliers when they exist and learn to investigate their source. Students learn that standard deviation is a measure of spread, that a larger standard deviation means the data are more spread out, and to understand standard deviation as "typical distance from the mean" for a symmetrical distribution. They also understand that interquartile range is a description of variability better-suited to a skewed distribution. Finally, students are introduced to two-way frequency tables and understand how to interpret relative frequencies in the context of the data represented in the tables.

M1.1.8 Summative Assessment

Assess students' ability to

• calculate mean, median, and mode;
• create box plots given data;
• compare and contrast two frequency distributions;
• articulate reasons to choose mean or use median as a measure of center;
• read and interpret relative frequencies.

Instead of creating representations of data, the emphasis in high school is on interpreting representations and judiciously interpreting measures of center and spread. Students describe the shape of a data distribution in more detail (symmetric, skewed, flat, or bell-shaped). Students develop a more precise understanding of measures of center and understand relationships between mean and median for symmetrical and skewed data distributions. They learn that outliers affect the mean of a data set but not the median. They recognize outliers when they exist and learn to investigate their source. Students learn that standard deviation is a measure of spread, that a larger standard deviation means the data are more spread out, and to understand standard deviation as "typical distance from the mean" for a symmetrical distribution. They also understand that interquartile range is a description of variability better-suited to a skewed distribution. Finally, students are introduced to two-way frequency tables and understand how to interpret relative frequencies in the context of the data represented in the tables.

M1.1.2 Analyze data distributions

• Create dot plots, histograms and box plots.
• Use available classroom technology to create histograms and box plots and calculate measures of center and spread.
• Use terms such as "flat," "skewed," "bell-shaped," and "symmetric" to describe data distributions.
• Analyze and compare data sets.

Instead of creating representations of data, the emphasis in high school is on interpreting representations and judiciously interpreting measures of center and spread. Students describe the shape of a data distribution in

more detail (symmetric, skewed, flat, or bell-shaped). Students develop a more precise understanding of measures of center and understand relationships between mean and median for symmetrical and skewed data distributions. They learn that outliers affect the mean of a data set but not the median. They recognize outliers when they exist and learn to investigate their source. Students learn that standard deviation is a measure of spread, that a larger standard deviation means the data are more spread out, and to understand standard deviation as "typical distance from the mean" for a symmetrical distribution. They also understand that interquartile range is a description of variability better-suited to a skewed distribution. Finally, students are introduced to two-way frequency tables and understand how to interpret relative frequencies in the context of the data represented in the tables.

M1.1.3 Measures of center

• Recall how to calculate mean and median.
• Understand mean and median as a "typical value" that can answer a statistical question.
• Know that mean and median are equal for a symmetrical data distribution.
• Explain why mean and median are unequal for a skewed data distribution.
• Select mean as the better measure for symmetrical distributions, and median as the better measure for skewed distributions.
• Make generalization what kinds of distributions have means larger than medians, and what kinds have medians larger than means.
• Recognize outliers when they exist, and know to investigate their source—that data point is way out there, why is that? Is there something weird about it that means we should disregard it?
• Know that outliers affect the mean, but not the median of a data set.

Instead of creating representations of data, the emphasis in high school is on interpreting representations and judiciously interpreting measures of center and spread. Students describe the shape of a data distribution in more detail (symmetric, skewed, flat, or bell-shaped). Students develop a more precise understanding of measures of center and understand relationships between mean and median for symmetrical and skewed data

distributions. They learn that outliers affect the mean of a data set but not the median. They recognize outliers when they exist and learn to investigate their source. Students learn that standard deviation is a measure of spread, that a larger standard deviation means the data are more spread out, and to understand standard deviation as "typical distance from the mean" for a symmetrical distribution. They also understand that interquartile range is a description of variability better-suited to a skewed distribution. Finally, students are introduced to two-way frequency tables and understand how to interpret relative frequencies in the context of the data represented in the tables.

M1.1.4 Mid-unit assessment

Assess students' ability to

• describe a set of data given a graph or table;
• identify and calculate spread, center, shape, outliers, quartiles, mean, median, mode;
• construct and interpret a box plot;
• compare, contrast, and draw conclusions when given two data sets.

Instead of creating representations of data, the emphasis in high school is on interpreting representations and judiciously interpreting measures of center and spread. Students describe the shape of a data distribution in more detail (symmetric, skewed, flat, or bell-shaped). Students develop a more precise understanding of measures of center and understand relationships between mean and median for symmetrical and skewed data distributions. They learn that outliers affect the mean of a data set but not the median. They recognize outliers when they exist and learn to investigate their source. Students learn that standard deviation is a measure of spread, that a larger standard deviation means the data are more spread out, and to understand standard deviation as "typical distance from the mean" for a symmetrical distribution. They also understand that interquartile range is a description of variability better-suited to a skewed distribution. Finally, students are introduced to two-way frequency tables and understand how to interpret relative frequencies in the context of the

data represented in the tables.

M1.1.5 Standard deviation

• Describe variability by calculating deviations from the mean.
• Compare two data sets with the same means but different variabilities, and contrast them by calculating the deviation of each data point from the mean.
• Interpret sets with greater deviations as having greater variability.
• Calculate a standard deviation by hand for a small data set, and understand standard deviation as an indicator of a typical deviation from the mean of an element of the data set.

Instead of creating representations of data, the emphasis in high school is on interpreting representations and judiciously interpreting measures of center and spread. Students describe the shape of a data distribution in more detail (symmetric, skewed, flat, or bell-shaped). Students develop a more precise understanding of measures of center and understand relationships between mean and median for symmetrical and skewed data distributions. They learn that outliers affect the mean of a data set but not the median. They recognize outliers when they exist and learn to investigate their source. Students learn that standard deviation is a measure of spread, that a larger standard deviation means the data are more spread out, and to understand standard deviation as "typical distance from the mean" for a symmetrical distribution. They also understand that interquartile range is a description of variability better-suited to a skewed distribution. Finally, students are introduced to two-way frequency tables and understand how to interpret relative frequencies in the context of the data represented in the tables.

M1.1.6 Bringing it all together

• Represent a data set in different ways and decide which way is most appropriate.
• Select measures of center and spread appropriate to the shape of the distribution.

• Compare and contrast two or more distributions by using appropriate measures to describe center, variability, and shape.

Instead of creating representations of data, the emphasis in high school is on interpreting representations and judiciously interpreting measures of center and spread. Students describe the shape of a data distribution in more detail (symmetric, skewed, flat, or bell-shaped). Students develop a more precise understanding of measures of center and understand relationships between mean and median for symmetrical and skewed data distributions. They learn that outliers affect the mean of a data set but not the median. They recognize outliers when they exist and learn to investigate their source. Students learn that standard deviation is a measure of spread, that a larger standard deviation means the data are more spread out, and to understand standard deviation as "typical distance from the mean" for a symmetrical distribution. They also understand that interquartile range is a description of variability better-suited to a skewed distribution. Finally, students are introduced to two-way frequency tables and understand how to interpret relative frequencies in the context of the data represented in the tables.

M1.1.7 Two-way frequency tables

• Interpret a two-way table.
• Understand that the choices made when organizing data can lead to different conclusions.

Instead of creating representations of data, the emphasis in high school is on interpreting representations and judiciously interpreting measures of center and spread. Students describe the shape of a data distribution in more detail (symmetric, skewed, flat, or bell-shaped). Students develop a more precise understanding of measures of center and understand relationships between mean and median for symmetrical and skewed data distributions. They learn that outliers affect the mean of a data set but not the median. They recognize outliers when they exist and learn to investigate their source. Students learn that standard deviation is a measure of spread, that a larger standard deviation means the data are

more spread out, and to understand standard deviation as “typical distance from the mean” for a symmetrical distribution. They also understand that interquartile range is a description of variability better-suited to a skewed distribution. Finally, students are introduced to two-way frequency tables and understand how to interpret relative frequencies in the context of the data represented in the tables.