HENRY COUNTY SCHOOLS

Updated 1/11/2013

# Unit 4
# Describing Data

## Key Vocabulary

- Association
- Bivariate data
- Box Plot
- Box and Whisker Plot
- Categorical Variables
- Center
- Conditional Frequencies
- Correlation Coefficient
- Dot Plot
- First Quartile
- Histogram
- Interquartile Range
- Joint Frequencies
- Line of Best Fit
- Marginal Frequencies
- Mean Absolute Deviation (MAD)
- Outlier
- Quantitative Variables
- Residuals
- Residual plot
- Scatter plot
- Second Quartile
- Shape
- Symmetry
- Number of Peaks
- Direction of Skew
- Uniformity
- Spread
- Third quartile
- Trend
- Two-Frequency Table

## CCGPS Standards Addressed:

MMCC9-12.S.ID.1
MMCC9-12.S.ID.2
MMCC9-12.S.ID.3
MMCC9-12.S.ID.5
MMCC9-12.S.ID.6a, b, c
MMCC9-12.S.ID.7
MMCC9-12.S.ID.8
MMCC9-12.S.ID.9

*Transition Standard*

MMCC6.SP.5c

## Prerequisite Skills

It is expected that students will have prior knowledge/experience related to the concepts and skills identified below. It may be necessary to pre-assess in order to determine if time needs to be spent on conceptual activities that help students develop a deeper understanding of these ideas.

- Know how to compute the mean, median, interquartile range, and mean standard deviation by hand in simple cases and using technology with larger data sets.
- Find the lower extreme (minimum), upper extreme (maximum), and quartiles.
- Create a graphical representation of a data set.
- Present data in a frequency table.
- Plot data on a coordinate grid and graph linear functions.
- Recognize characteristics of linear and exponential functions.
- Write an equation of a line given two points.
- Graph data in a scatter plot and determine a trend.
- Determine the slope of a line from any representation.
- Identify the y-intercept from any representation.
- Be able to use graphing technology.
- Understand the meaning of correlation.

## Essential Question

*How do I summarize, represent, and interpret data on a single count or measurement variable?*

*When taking real-life actions, what factors are important for me to consider in determining which statistics to compare, graphical representation, and interpretation techniques?*

*How do I summarize, represent, and interpret data on two categorical and quantitative variables?*

*Why is technology valuable when making statistical models?*

*Why are linear models used to study many important real-world phenomena?*

*How do I interpret linear models?*

## Suggested Learning Resources/ Performance Tasks

- Math Task (Closing in Homework Task)
- If the Shoe Fits
- Spaghetti Regressions
- The Basketball Star
- Public Opinions
- Leisure Time
- TV Test Grades
- Equal Salaries for Equal Work
- Research Design

## Enduring Understandings

Students will understand…

- Data are gathered, displayed, summarized, examined, and interpreted to discover patterns and deviations from patterns.
- Which statistics to compare, which plots to use, and what the results of a comparison might mean, depend on the question to be investigated and the real-life actions to be taken.
- Understand and be able to use the context of the data to explain why its distribution takes on a particular shape (e.g. are there real-life limits to the values of the data that force skewness?).
- When making statistical models, technology is valuable for varying assumptions, exploring consequences and comparing predictions with data.
- Causation implies correlation yet correlation does not imply causation.

## Key CCGPS Standards Addressed:

~~TRANSITION STANDARDS FOR 2012-13 ONLY~~

~~MCC9-12.S.ID.4 Summarize a single data set in relation to its context, such as by:~~
~~a. Giving quantitative measures of center (median and/or mean) and variability (interquartile range and/or mean absolute deviation), as well as describing any overall pattern and any striking deviations from the overall pattern with reference to the context in which the data was gathered.~~

**MCC9-12.S.ID.1** Represent data with plots on the real number line (dot plots, histograms, and box plots). Choose appropriate graphs to be consistent with numerical data: dot plots, histograms, and box plots.

**MCC9-12.S.ID.2** Use statistics appropriate to the shape of the data distribution to compare center (median, mean) and spread (interquartile range, standard deviation-Advanced Algebra) of two or more different data sets. Include review of Mean Absolute Deviation as a measure of variation.

**MCC9-12.S.ID.3** Interpret differences in shape, center, and spread in the context of the data sets, accounting for possible effects of extreme data points (outliers). Students will examine graphical representations to determine if data are symmetric, skewed left, or skewed right and how the shape of the data affects descriptive statistics.

**MCC9-12.S.ID.5** Summarize categorical data for two categories in two-way frequency tables. Interpret relative frequencies in the context of the data (including joint, marginal, and conditional relative frequencies). Recognize possible associations and trends in the data.

**MCC9-12.S.ID.6** Represent data on two quantitative variables on a scatter plot, and describe how the variables are related.

**MCC9-12.S.ID.6a** Fit a function to the data; use functions fitted to data to solve problems in the context of the data. Use given functions or chooses a function suggested by the context. Emphasize linear, quadratic, and exponential models.

**MCC9-12.S.ID.6b** Informally assess the fit of a function by plotting and analyzing residuals.

**MCC9-12.S.ID.6c** Fit a linear function for a scatter plot that suggests a linear association.

**MCC9-12.S.ID.7** Interpret the slope (rate of change) and the intercept (constant term) of a linear model in the context of the data.

**MCC9-12.S.ID.8** Compute (using technology) and interpret the correlation coefficient of a linear fit.

**MCC9-12.S.ID.9** Distinguish between correlation and causation.