M1.3 Bivariate Statistics

• Represent data on two quantitative variables on a scatter plot.
• Describe how two quantitative variables on a scatter plot are related.
• Interpret the slope and the intercept of a linear model in the context of the data.
• Use available technology to find lines of best fit.
• Assess the goodness of fit of a line to a small data set by plotting and analyzing residuals.
• Fit a linear function for a scatter plot that suggests a linear association.
• Use available technology to compute correlation coefficients.
• Understand that the correlation coefficient measures the "tightness" of a line fitted to data.
• Understand that correlation does not necessarily imply causality.

In this unit, students build on the statistical work they did in grade 8. Using technology, they represent data on two variables on a scatter plot, find the line of best fit, and interpret the slope and intercepts in context. They assess the fit of a line to data more precisely by plotting and analyzing residuals. Students compare strength of associations between different pairs of variables by interpreting correlation coefficients which they compute using technology. They gain experience in distinguishing between correlation and causality. Modeling is an intrinsic part of the high school statistics and probability standards and of this unit.

From their experiences with linear functions in grade 8, students are familiar with slope and intercept. They gained experience with scatter plots and described patterns such as clustering, outliers, positive or negative association, linear association, and nonlinear association. (8.SP.A.1) For scatter plots that suggested

a linear association, they informally fit a line and informally assessed its fit. (8.SP.A.2) They wrote equations for these linear models, and interpreted their slopes and intercepts in the context of the data. (8.SP.A.3) In this unit, students build on the statistical work they did in grade 8. They work with bivariate data and find the line of best fit by using a graphing calculator or other software. (These lines of best fit are regression lines (or technologically-generated approximations of them) but the Standards do not require students to learn or use the terms "regression," "regression line," "regression equation," or "least squares.") They assess the fit of a line to data more precisely by plotting and analyzing residuals. Students compare strength of associations between different pairs of variables by interpreting correlation coefficients, which they compute using technology. And, they gain experience in distinguishing between correlation and causality. Modeling is an intrinsic part of the high school statistics and probability standards, and of this unit. In unit S4, student build on the techniques used this unit. Plotting residuals suggests that not some data sets may be better modeled non-linear functions than linear ones.

M1.3.0 Pre-unit diagnostic assessment

Assess students' ability to

• write a linear equation and interpret it in a context;
• determine a rate of change and initial value given several data points that exhibit a linear relationship;
• write an equation for a linear relationship;
• use the equation to make predictions;
• interpret the slope and intercept of a linear equation in a context.

In this unit, students build on the statistical work they did in grade 8. Using technology, they represent data on two variables on a scatter plot, find the line of best fit, and interpret the slope and intercepts in context. They assess the fit of a line to data more precisely by plotting and analyzing residuals. Students compare strength of associations between different pairs of variables by interpreting correlation coefficients which they compute using technology. They gain experience in distinguishing between correlation and causality. Modeling is an intrinsic part of the high school statistics and

probability standards and of this unit.

M1.3.1 Preview

• Activate prior experience in grade 8 with informally fitting a line to a scatter plot and informally judging its goodness of fit.
• Model the relationship with an equation for a line and use it to make predictions.
• Represent data on two quantitative variables on a scatter plot.
• Describe how two variables on a scatter plot are related.
• Interpret the slope and the intercept of a linear model in the context of the data.

(Note: 8.SP.2 and 8.SP.3 are prerequisites, not target standards in this unit. However, they are standards involved in one of the suggested activities.)

In this unit, students build on the statistical work they did in grade 8. Using technology, they represent data on two variables on a scatter plot, find the line of best fit, and interpret the slope and intercepts in context. They assess the fit of a line to data more precisely by plotting and analyzing residuals. Students compare strength of associations between different pairs of variables by interpreting correlation coefficients which they compute using technology. They gain experience in distinguishing between correlation and causality. Modeling is an intrinsic part of the high school statistics and probability standards and of this unit.

M1.3.2 Lines of best fit and residuals

• Use available technology to find lines of best fit.
• Quantify the goodness of fit by plotting and analyzing residuals.

In this unit, students build on the statistical work they did in grade 8. Using technology, they represent data on two variables on a scatter plot, find the line of best fit, and interpret the slope and intercepts in context. They assess the fit of a line to data more precisely by plotting and analyzing residuals. Students compare strength of associations between different pairs of

variables by interpreting correlation coefficients which they compute using technology. They gain experience in distinguishing between correlation and causality. Modeling is an intrinsic part of the high school statistics and probability standards and of this unit.

M1.3.3 Interpreting the correlation coefficient

• Fit a linear function for a scatter plot that suggests a linear association.
• Use available technology to compute correlation coefficients.
• Understand that the correlation coefficient measures the "tightness" of a line fitted to data.
• Understand the significance of correlation coefficients close to 1 or –1.
• Interpret the rate of change and constant term of a line fitted to data in the context of the data.

In this unit, students build on the statistical work they did in grade 8. Using technology, they represent data on two variables on a scatter plot, find the line of best fit, and interpret the slope and intercepts in context. They assess the fit of a line to data more precisely by plotting and analyzing residuals. Students compare strength of associations between different pairs of variables by interpreting correlation coefficients which they compute using technology. They gain experience in distinguishing between correlation and causality. Modeling is an intrinsic part of the high school statistics and probability standards and of this unit.

M1.3.4 Correlation vs causation

Understand that correlation does not necessarily imply causality.

In this unit, students build on the statistical work they did in grade 8. Using technology, they represent data on two variables on a scatter plot, find the line of best fit, and interpret the slope and intercepts in context. They assess the fit of a line to data more precisely by plotting and analyzing residuals. Students compare strength of associations between different pairs of variables by interpreting correlation coefficients which they compute using technology. They gain experience in distinguishing between correlation and

causality. Modeling is an intrinsic part of the high school statistics and probability standards and of this unit.

M1.3.5 Bringing it all together

• Describe how two quantitative variables are related.
• Use technology to create a line of best fit.
• Fit a line to given data and use it to make predictions.
• Interpret the coefficients of a line of best fit in the context of the data to which it is fitted.
• Compute and interpret a correlation coefficient.
• Understand that variables with a high correlation do not necessarily have a causal relationship.

In this unit, students build on the statistical work they did in grade 8. Using technology, they represent data on two variables on a scatter plot, find the line of best fit, and interpret the slope and intercepts in context. They assess the fit of a line to data more precisely by plotting and analyzing residuals. Students compare strength of associations between different pairs of variables by interpreting correlation coefficients which they compute using technology. They gain experience in distinguishing between correlation and causality. Modeling is an intrinsic part of the high school statistics and probability standards and of this unit.

M1.3.6 Summative Assessment

Assess students' ability to

• create a scatter plot and a line of best fit using technology;
• interpret the coefficients of the line of best fit in the context of the data;
• interpret the correlation coefficient;
• articulate the difference between correlation and causation.

In this unit, students build on the statistical work they did in grade 8. Using technology, they represent data on two variables on a scatter plot, find the line of best fit, and interpret the slope and intercepts in context. They assess

the fit of a line to data more precisely by plotting and analyzing residuals. Students compare strength of associations between different pairs of variables by interpreting correlation coefficients which they compute using technology. They gain experience in distinguishing between correlation and causality. Modeling is an intrinsic part of the high school statistics and probability standards and of this unit.