MATH 2

UNIT 6

Linear and Quadratic Regression

Georgia Performance Standards
High School Mathematics
Mathematics 2

Georgia Performance Standards: Curriculum Map					
1st Semester			2nd Semester		
Unit 1	Unit 2	Unit 3	Unit 4	Unit 5	Unit 6
Quadratics & Complex Numbers	Right Triangle Trigonometry	Circles & Spheres	Data Analysis & Probability	Step and Piecewise Functions	Linear and Quadratic Regression
6 Weeks	4 Weeks	6 Weeks	5 Weeks	6 Weeks	4 Weeks
MM2N1 MM2A3 MM2A4	MM2G1 MM2G2	MM2G3 MM2G4	MM2D1	MM2A1b MM2A2 MM2A5	MM2A1a, c MM2D2

These units were written to build upon concepts from prior units, so later units contain tasks that depend upon the concepts addressed in earlier units. Standards listed are key standards for the units. All units will include the Process Standards.

MATH 2 UNIT 6 CONTENT MAP
Finding the Best Model

Unit 6 – Finding the Best Model (4 Weeks)

Essential Questions:

Lesson 1 – Absolute Value Functions (2 hours)
Essential Question: How do you analyze and graph absolute value functions?

Lesson 2 – Solving Absolute Value Equations and Inequalities (5 hours)
Essential Question: How do you solve absolute value equations and inequalities algebraically and graphically?

Lesson 3 – Line of Best Fit (5 hours)
Essential Question: How do you determine the line of best fit using linear and quadratic regression for a set of data points using various methods such as eyeballing, median-median line, and technology?

Lesson 4 – Correlation and Causation: (5 Hours)
Essential Question: How do you explore the relationship between variables?

Summarization, Review, & Evaluation of Unit 1 (3 Hours)

Mathematics 2 – Unit 6: Linear and Quadratic Regression

INTRODUCTION:

This is the final unit of Mathematics II. Students learn to use statistical tools to find functions that are good mathematical models of sets of data. Much of the computation is done with the use of appropriate technology so that the focus of the unit is understanding when and how to use each tool and determining when a function is a "best fit" for a set of data. In the process students revisit important characteristics of linear and quadratic functions. Students also extend their knowledge of absolute value functions by constructing an absolute value function to model a real-world situation. As they work with this model, they extend their understanding of absolute value functions as piecewise functions and learn to solve absolute value equations and inequalities.

ENDURING UNDERSTANDINGS:

- Absolute value equations and inequalities can be solved both algebraically and graphically.
- Choosing an appropriate model for a set of data requires examining a plot of the data and analyzing the fit of the data to the model.
- Correlation provides information about the strength and direction of a linear relationship, but it does not indicate causation between the two quantitative variables.
- Most data do not fit any one function rule exactly; some amount of error usually exists.
- The method of finite differences is a tool for deciding if a set of data exactly fits a linear, quadratic, or other polynomial model.

KEY STANDARDS ADDRESSED:

MM2A1. Students will investigate step and piecewise functions, including greatest integer and absolute value functions.

a. Write absolute value functions as piecewise functions.

c. Solve absolute value equations and inequalities analytically, graphically, and by using appropriate technology.

MM2D2. Students will determine an algebraic model to quantify the association between two quantitative variables.

a. Gather and plot data that can be modeled with linear and quadratic functions.

b. Examine the issues of curve fitting by finding good linear fits to data using simple methods such as the median-median line and "eyeballing."

c. Understand and apply the processes of linear and quadratic regression for curve fitting using appropriate technology.

d. Investigate issues that arise when using data to explore the relationship between two variables, including confusion between correlation and causation.

RELATED STANDARDS ADDRESSED:

MM2P1. Students will solve problems (using appropriate technology).

a. Build new mathematical knowledge through problem solving.

b. Solve problems that arise in mathematics and in other contexts.

c. Apply and adapt a variety of appropriate strategies to solve problems.

d. Monitor and reflect on the process of mathematical problem solving.

MM2P2. Students will reason and evaluate mathematical arguments.

a. Recognize reasoning and proof as fundamental aspects of mathematics.
b. Make and investigate mathematical conjectures.
c. Develop and evaluate mathematical arguments and proofs.
d. Select and use various types of reasoning and methods of proof.

MM2P3. Students will communicate mathematically.

a. Organize and consolidate their mathematical thinking through communication.
b. Communicate their mathematical thinking coherently and clearly to peers, teachers, and others.
c. Analyze and evaluate the mathematical thinking and strategies of others.
d. Use the language of mathematics to express mathematical ideas precisely.

MM2P4. Students will make connections among mathematical ideas and to other disciplines.

a. Recognize and use connections among mathematical ideas.
b. Understand how mathematical ideas interconnect and build on one another to produce a coherent whole.
c. Recognize and apply mathematics in contexts outside of mathematics.

MM2P5. Students will represent mathematics in multiple ways.

a. Create and use representations to organize, record, and communicate mathematical ideas.
b. Select, apply, and translate among mathematical representations to solve problems.
c. Use representations to model and interpret physical, social, and mathematical phenomena.

UNIT OVERVIEW:

The unit begins with the real-world context of a schedule of tolls from the Ohio turnpike. This context is used for an extensive two-part task. The first part uses turnpike tolls as data sets for which students find lines of best fit by several methods, visual approximation, median-median lines, and least squares linear regression. The second part of the task involves creating of an absolute value function to model the part of the toll schedule relating to one particular exit. This activity involves further exploration of absolute value functions as piecewise functions and provides motivation for solving the absolute values equations and inequalities, both algebraically and with the use of graphing technology.

The second learning task reviews least squares linear regression, introduces correlation coefficient, and explores least squares quadratic regression through the context of studying the orbital debris in space.

The third task requires students to do some research to find specific examples to illustrate that a strong correlation, as measure by correlation coefficient, does not mean that one of phenomenon causes another.

The last task has students work in pairs to gather data about the diameter and area of circles visible when viewed through a tube at various distances from a wall. Students explore appropriate models for the data and learn about the method of finite differences for determining whether a set of data exactly fit a linear, quadratic, or other polynomial model. In

working with real data that they have collected through measurement, they further explore how real data that is well-modeled by a linear or quadratic function differs from an exact linear or quadratic model.

The overall theme of the unit is finding linear, quadratic, or absolute value functions to model sets of data. A major emphasis is placed on analyzing models for goodness of fit and understanding issues in using the models to predict other data points for the same context. Throughout this unit it is assumed that students have access to a graphing utility that can graph scatter plots and graphs of functions given by formula on the same coordinate axes and to technology for calculating median-median lines, least squares regression lines, correlation coefficients, and quadratic regression.

Throughout the unit, there needs to be careful attention to use of correct terminology in reference to data sets, statistical computations, and functions which model sets of data. The unit also involves analysis of many statistical issues that require discussion and explanation. Therefore, throughout this unit, students should regularly communicate their reasoning about the mathematics and the statistical issues in writing as well as orally in class discussions.

Vocabulary and Formulas:

Correlation coefficient – The correlation coefficient, *r*, measures the direction and strength of a *linear* relationship between two variables. Formula: $r = \frac{1}{n-1}\sum\left(\frac{x_i - \bar{x}}{s_x}\right)\left(\frac{y_i - \bar{y}}{s_y}\right)$.

Extrapolation – the use of a regression curve to make predictions for a value of the independent variable less than the smallest, or greater than the largest, value of the independent variable occurring with the data set that the regression curve models.

Interpolation – the use of a regression curve to make predictions for a value of the independent variable that is between two values of the independent variable occurring with the data set that the regression curve models.

Linear regression line – A straight line that approximates the relationship between two variables represented by a set of data points.

Least squares regression line (LSRL) – the line that minimizes the sum of the squares of the vertical distances between the data points and any possible regression line.

Median-median line – a linear regression line found by a method based on the calculation of medians. This method of linear regression requires that the data points are ordered from smallest to largest first coordinate and then separates the data into three equal, or nearly equal, groups with at least 1/3 of the data points in each of the first and last groups. The median x-values and y-values of each group are calculated. These medians, from smallest x-values to largest, are named $(x_1, y_1), (x_2, y_2), (x_3, y_3)$. Then a line through the first and third medians is found. Finally, a line parallel to this line, 1/3 of the distance between the line and the remaining median is formed. The resulting line is of the following form $y = ax + b, a = \frac{y_3 - y_1}{x_3 - x_1}, b = \frac{y_1 + y_2 + y_3 - a(x_1 + x_2 + x_3)}{3}$. This method of regression is more resistant to outliers than the least squares regression line.

Method of finite differences – a method for determining if data points with equally spaced *x*-values exactly fit a linear, quadratic, or higher degree polynomial model.

Quadratic Regression – a quadratic function that minimizes the sum of the squares of the vertical distances between the data points and any possible quadratic function to approximate the data.

Regression curve – the graph of a function, including possibly a linear function, that approximates the relationship between two variables represented by a set of data points. (Linear and quadratic regression are explored in this unit.)

TASKS:

RESOURCES NEEDED BY THE TEACHER FOR THE LESSONS IN THIS UNIT:
Elmo or Overhead Projector for each teacher, Classroom set of Graphing Calculators, Classroom set of Algebra Tiles, Classroom set of small cubes such as Algeblocks, Classroom set of individual marker boards (blank on one side and a grid on the other) and markers for students to use, Coordinate Grids, Colored Pencils, Rulers, Markers, Roll of Graph Paper with Inch Squares, Pad of Quad Paper, Glue Sticks, Scissors, Post-it Notes, Construction Paper, Poster Board, Copies of all Handouts for Students, Copies of the Standards for Students, Large Copy of the Standards to Post on the Wall

RESOURCES NEEDED BY THE STUDENTS FOR THE LESSONS IN THIS UNIT:
Notebook with at least 10 dividers for the introduction, individual lessons, and culminating activities, pencils, notebook paper, graph paper

Note: A copy of the standards for this unit should be given to the students with discussion to be held throughout the unit concerning their meaning and relation to the learning tasks of the day. A large copy of the standards should be posted in the classroom and referred to regularly as the various standards are discussed. Vocabulary should be emphasized by using word walls and flash cards or vocabulary sheets in the students' notebooks. Students will need individual copies of all handouts in the lessons of the unit. Students should keep a math notebook with all materials in it.