Unit 4: Multivariate Data Analysis

Content Area: **Mathematics**
Course(s):
Time Period: **Marking Period 2**
Length: **6 weeks**
Status: **Published**

Summary of Multivariate Data Analysis

As an extension of the previous unit, students are now presented with tools of data analysis to explore interactions among multiple variables. We first look at bivariate relationships between independent (explanatory) variables and dependent (response) variables, and examine the association for correlation or causation. If an association exists, it may be measured, expressed mathematically, and used to predict new values. Linear and exponential models will be examined, with greater emphasis on the construction and interpretation of linear regression models. Given that most real life situations involve more than just two factors, this unit also explores situations that are mathematically modeled by multiple variables. Students will learn the process of how massive amounts of data comprising many different variables are funneled into a single number, called an index. These indexes are used to compare complex data across various subjects and over time. Multivariate indexes are statistical measures used in many academic fields; students will learn how to construct and interpret a variety of widely-used indexes, and will also work collaboratively to create original indexes related to their own areas of interest. This unit includes an examination of the Consumer Price Index, combined with cost of living adjustments and historical records of minimum wage, to inform students with a better understanding of the U.S. economy as it relates to real earnings and matters of affordability, including the cost of college education over the last seven decades.

Revision Date: July 2021

Analyzing various sets of data will allow students to explore studies about people from different backgrounds. Statistical studies and analysis provides students an opportunity to read about historical statistics about people's cultures. Embracing the diversity within society incorporates the following:

Amistad Commission

This unit also reflects the goals of the Department of Education and the Amistad Commission including the infusion of the history of Africans and African-Americans into the curriculum in order to provide an accurate, complete, and inclusive history regarding the importance of of African-Americans to the growth and development of American society in a global context.

Asian American and Pacific Islander History Law

This unit includes instructional materials that highlight the history and contributions of Asian Americans and Pacific Islanders in accordance with the New Jersey Student Learning Standards

in Social Studies.

New Jersey Diversity and Inclusion Law

In accordance with New Jersey's Chapter 32 Diversity and Inclusion Law, this unit includes instructional materials that highlight and promote diversity, including:

economic diversity, equity, inclusion, tolerance, and belonging in connection with gender and sexual orientation, race and ethnicity, disabilities, and religious tolerance.

Commission on Holocaust Education

This unit further reflects the goals of the Holocaut Education mandate where students are able to identify and analyze applicable theories concerning human nature and behavior; understand that genocide is a consequence of prejudice and discrimination; understand that issues of moral dilemma and conscience have a profound impact on life; and understand the personal responsibility that each citizen bears to fight racism and hatred whenever and wherever it happens.

MA.S-ID.A.2	Use statistics appropriate to the shape of the data distribution to compare center (median, mean) and spread (interquartile range, standard deviation) of two or more different data sets.
MA.S-ID.B	Summarize, represent, and interpret data on two categorical and quantitative variables
MA.S-ID.B.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.
MA.S-ID.B.6	Represent data on two quantitative variables on a scatter plot, and describe how the variables are related.
MA.S-ID.B.6a	Fit a function to the data (including with the use of technology); use functions fitted to data to solve problems in the context of the data.
MA.S-ID.B.6c	Fit a linear function for a scatter plot that suggests a linear association.
MA.S-ID.C	Interpret linear models
MA.S-ID.C.7	Interpret the slope (rate of change) and the intercept (constant term) of a linear model in the context of the data.
MA.S-ID.C.9	Distinguish between correlation and causation.
MA.S-IC.B	Make inferences and justify conclusions from sample surveys, experiments, and observational studies
LA.K-12.NJSLSA.W4	Produce clear and coherent writing in which the development, organization, and style are appropriate to task, purpose, and audience.
LA.K-12.NJSLSA.W5	Develop and strengthen writing as needed by planning, revising, editing, rewriting, or trying a new approach.

MA.S-IC.B.6	Evaluate reports based on data.
WRK.K-12.P.5	Utilize critical thinking to make sense of problems and persevere in solving them.
TECH.9.4.12.CI.1	Demonstrate the ability to reflect, analyze, and use creative skills and ideas (e.g., 1.1.12prof.CR3a).
TECH.9.4.12.CT.2	Explain the potential benefits of collaborating to enhance critical thinking and problem solving (e.g., 1.3E.12profCR3.a).
TECH.9.4.12.IML.3	Analyze data using tools and models to make valid and reliable claims, or to determine optimal design solutions (e.g., S-ID.B.6a., 8.1.12.DA.5, 7.1.IH.IPRET.8).
TECH.9.4.12.IML.4	Assess and critique the appropriateness and impact of existing data visualizations for an intended audience (e.g., S-ID.B.6b, HS-LS2-4).

Essential Questions / Enduring Understandings

- What is the difference between correlation and causation?
- How can the tools of algebra be used to create linear models to better understand relationships among bivariate data sets?
- What processes allow us to create mathematical models for multivariate data?
- How can mathematical models of multivariate data be used to make comparisons across subjects and over time?
- What role does the Consumer Price Index (CPI) play in providing insight to the U.S. economy?
- Linear regression models can be used to understand bivariate relationships and predict values for both explanatory and response variables.
- Independent and dependent variables may have positive or negative associations, or no associations at all.
- Correlation does not imply causation.
- A multivariate index is a statistical measure that combines data from a number of sources to produce a single quantitative value used to interpret and compare complex situations that are affected by multiple variables.
- Indexes are constructed using a variety of mathematical operations, may incorporate different categorical weights, and are used within many academic fields.

Objectives for Multivariate Data Analysis

Students Will Know

- How to construct and interpret mathematical models for bivariate data.
- How to use linear regression to predict values of independent and dependent variables.
- Confounding variables may exist and lead to false conclusions.
- Multivariate indexes incorporate a variety of mathematical processes with multiple sources of data.
- The role of indexes for data analysis and comparative studies across complex situations with many variables.

Objectives for Multivariate Data Analysis

Students Will Be Skilled At

- Creating linear models of bivariate data sets, and recognizing nonlinear relationships.
- Working with bivariate and multivariate mathematical models to understand relationships among variables and make predictions for values which lie outside the known data sets.
- Using multivariate indexes for comparative analyses.
- Using real data to calculate and interpret some of the following widely-recognized indexes:
- Weather indexes, including heat index and wind chill
- Academic indexes, including the fog index and the college student stress index
- Sociopolitical indexes, including the shoe-thrower's index, the misery index, human development index, global hunger index, and the world peace index
- Consumer Price Index (CPI) in conjunction with Cost of Living Adjustments (COLA) to better understand socioeconomic portrayals of life in the U.S.A.

Learning Plan for Multivariate Data Analysis

The topic of linear models for bivariate relationships is presented earlier in the sequence of secondary mathematics as a study of scatter plots in traditional algebra curricula. In this unit, students will be presented with additional tools to further analyze bivariate data and create models for interpretative and predictive purposes. To begin this unit, the instructor will review fundamental concepts from algebra about linear graphs on the coordinate plane, including the identification, calculation, and interpretation of the slope and intercepts. Students will review the equation of a line, and apply this prior knowledge to the practice of constructing a line of best fit for bivariate data sets. The instructor will use class time to present this content, clarify students' understanding of fundamental concepts, and provide many classroom opportunities to strengthen their skills with these matters. Short entrance and exit tickets will serve to formatively assess students' ability with constructing graphs and algebraic models, and will be followed by brief whole class discussions to ascertain correct answers. Unlike with bivariate data, it is doubtful that students have had prior experience with the mathematics of indexes. Accordingly, the instructor needs to preview the essential questions and introduce the new ideas surrounding multivariate data analysis with many real-life examples and explanations. Students should participate in guided note-taking during class time and carefully read through corresponding textbook sections and online handouts provided by the teacher. In this unit of study, the instructor will present many sets of practice exercises for students to complete while working together in small groups. Given the nature and unfamiliarity of the topic, it is important for students to collaborate on their explorations and interpretations of multivariate indexes during class time.

Evidence/Performance Tasks for Multivariate Data Analysis

Formative assessments will include classroom activities and discussions, with immediate feedback from the instructor. Short written quizzes will be administered and reviewed during class time in order to identify any challenges that students may have with their comprehension of these topics. Student mastery of bivariate data

analysis will be assessed through their ability to successfully construct linear graphs and algebraic models of original data sets presented on a written test during class time. With the multivariate data, students will have opportunities to demonstrate their understanding and will receive feedback with their work on computing and interpreting values with multivariate indexes. In preparation for an original project, students will study a collection of real indexes used by data analysts around the world for comparative and predictive purposes, and they will learn how a variety of mathematical operations may be combined to effectively produce a single value that can be understood and communicated to other people. Students will collaborate in this unit to create an original index based on personal and professional interests. That is, they will choose small groups of their peers who share similar interests, identify critical categories of data, research and gather real data, and compile an original multivariate model to be used for interpretations and comparisons of various subjects. Student groups will present their findings, respond to feedback from their classmates, and make necessary adjustments to their final index formula. Students' original projects will serve as both formative and summative assessments due to the requirement that students will necessarily meet with the instructor at designated checkpoints along the way. Additionally, students may have the option to verbally articulate their understanding of key concepts or elaborate on their written work. In this unit, students will be evaluated on their project work and the presentation of their findings to their peers.

Assessments

- Formative: Daily assessments using examples from class notes, NJSLA test bank problems, and/or Albert/AP Classroom assessments
- Summative: Teacher-created assessments, NJSLA test bank problems, Big Ideas Math online platform problems, Albert/AP Classroom and/or Big Ideas Math unit assessments
- Benchmark: IXL or teacher created diagnostic assessments in addition to unit assessments
- Alternative Assessments: Student-centered activities such as scavenger hunts, various projects involving real world applications, and differentiated learning tasks in Khan Academy, DeltaMath, and IXL

Materials for the Multivariate Data Analysis

Core instructional materials: Core Book List including Mathematical Ideas 14E, Miller, Heeren, & Hornsby, Savvas

Supplemental materials:

Lecture notes and classroom activities designed by instructor

Internet resources, including Khan Academy, YouTube instructional videos and teacher-recommended math

education websites