Unit 3: Univariate Data Analysis

Content Area: **Mathematics**
Course(s):
Time Period: **Marking Period 2**
Length: **4 weeks**
Status: **Published**

Summary of Univariate Data Analysis

Center and spread are the two key descriptors for a set of univariate data. As a measure of central tendency, we look at the mean, median, and mode. With regards to the spread of the univariate set, we look at the overall distribution of the data, giving attention to the shape, clusters, gaps, and outliers. The range of a data set tells us how spread out the maximum and the minimum values are, and the interquartile range provides insight into how the middle 50% of the values are distributed. The standard deviation is an advanced statistical calculation that describes the variation in the data set by looking at the relationship between each data value and the overall sample mean; standard deviation is the average distance of each data point from the mean. Upon studying a set of quantitative data, it is useful to create a visual representation or graph of the data to help with its interpretation. Exploratory analysis of data begins with visual representations of data. There are four common graphical representations that we use for univariate data, including a histogram, stem-and-leaf plot, dot plot, and box-and-whisker plot. In this unit, students will be trained to observe patterns, determine relationships, and draw conclusions about the data.

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	Summarize, represent, and interpret data on a single count or measurement variable
MA.S-ID.A.1	Represent data with plots on the real number line (dot plots, histograms, and box plots).
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.A.3	Interpret differences in shape, center, and spread in the context of the data sets, accounting for possible effects of extreme data points (outliers).
MA.S-ID.A.4	Use the mean and standard deviation of a data set to fit it to a normal distribution and to estimate population percentages. Recognize that there are data sets for which such a procedure is not appropriate. Use calculators, spreadsheets, and tables to estimate areas under the normal curve.
MA.S-IC.A	Understand and evaluate random processes underlying statistical experiments
MA.S-IC.A.1	Understand statistics as a process for making inferences about population parameters based on a random sample from that population.
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.
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	Critical Thinking and Problem-solving
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 are the measures of central tendency for a data set?
- How may one describe the amount of variation among values within a data set?
- How can describing the center, shape, and spread of a data set be useful in drawing conclusions about a population?
- How is standard deviation computed, and what does it reveal about a data set that is normally distributed?
- What is a z-score and how is it used to interpret the position of individual values, relative to the rest of the data values in the distribution?
- Values for the mean, median, and mode serve as measures of central tendency for a univariate data set.
- Outliers are extreme values that fall far outside the distribution of the data values and are quantitatively identified through a statistical measure that uses the spread of the interquartile range.
- The standard deviation is a critical value for use with the interpretation of a data set that is normally distributed.
- Visual representations provide quick insight with the data and are useful tools during initial exploratory phase of analysis.
- When represented with a histogram, data within a normal distribution fall into a symmetrical pattern that resembles a bell curve.

Objectives for Univariate Data Analysis

Students Will Know

- How to interpret visual representations of data.
- How to calculate the standard deviation for a data set.
- How to find corresponding percentiles for specific data values.
- How to describe the center, shape, and spread of a data set, and use those descriptors to help draw conclusions about a population.
- To identify the mean, median, and mode as measures of central tendency and indicators of typical values for a data set.
- Students will be able to describe the variation of data and know that the variation is as important as defining the center of a data set.
- The relation between a sample statistic and the probability of an event.
- How to analyze data from an original research project to infer a conclusion.

Objectives for Univariate Data Analysis

Students Will Be Skilled At

- Identifying the mean, median, and mode of a data set to describe its central tendency.
- Using different measures of spread, including range, interquartile range, and standard deviation to describe the spread of a data set.
- Constructing dot plots and stem-and-leaf plots.
- Using the five-number summary create a box-and-whisker plot.
- Using a quantitative measure to identify the presence of an outlier.
- Using the empirical rule for normal distributions; applying the mean and standard deviation to highlight important values in the data set.
- Utilizing the z-score formula to identify positions and percentile ranks for particular values in a normal distribution.
- Sorting and organizing raw data into tables for analysis.

Learning Plan for Univariate Data Analysis

This unit should begin by referencing students' ongoing research projects with the real data they have gathered. In this part of the course, students work with exploratory tools of data analysis; the motivation coming from their interest in better understanding their own individual research topics and trying to make sense of their unique data sets. During instructional class time, the teacher will present content and clarify concepts, and respond to students' questions. The students will be expected to read the assigned material, as well as supplementary resources found online. The tools of univariate data analysis will be presented in the context of real-life examples, and students will engage in guided note-taking activities to practice the formulas and learn the mechanics of the various types of graphical representations presented in this unit. Classroom time will be used for students to collaborate on exercises that develop and reinforce their understanding of exploratory analysis, and the instructor will facilitate whole group discussions about the topics. Throughout the unit, the students should work with real data, in relevant and engaging contexts. Accordingly, the instructor is encouraged to solicit input from the students about particular areas of interest, and to acquire appropriate data sets that correspond with their interests.

Evidence/Performance Tasks for Univariate Data Analysis

To develop skills with univariate data analysis, students will need to repetitively practice the tasks of identifying and calculating key numerical descriptors, as well as constructing the various types of visual representations for data of one variable. Formative assessments will include written tasks which ask students to calculate values that measure the center and spread of data sets. The teacher may assign these formative assessments as either entrance or exit tickets, and then utilize the class time that follows to provide immediate feedback about the answers. In addition, daily classroom activities and discussions will serve as formative assessments with the teacher sharing feedback and commentary at the individual and whole group levels. Students will also have opportunities to exchange input with their peers during small group exercises which provide opportunities for them to practice working with the content. As a summative assessment, all

students will conduct a thorough analysis on the univariate data they gathered for their original research projects. Students will demonstrate their comprehension of quantitative tools, as well as their ability to construct meaningful graphs of their data, and combine both forms of analysis in a presentation 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 Univariate 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