Data and Analysis

Design Brief for Standards Writers

SUGGESTED CITATION: Computer Science Teachers Association. (2026). *Data and analysis: Design brief for standards writers*. https://csteachers.org/k12standards/research

This design brief focuses on the following learning outcomes for the Data and Analysis topic area described in the *Reimagining CS Pathways* report:

- **DA.1** Identify and define data types (e.g., string, numeric, Boolean)
- **DA.6** Transform and prepare (e.g., normalize, merge, clean) data
- **DA.10** Evaluate approaches to cleaning data in a given context
- **DA.11** Assess whether and how a given question can be answered using computational methods and data, and what specific data is needed
- **DA.13** Evaluate data visualizations for clarity, potential biases, etc.
- **DA.14** Select, organize, interpret, and visualize large datasets from multiple sources to support a claim and/or communicate information
- **DA.15** Devise plans for using data to solve a problem
- **DA.16** Create a data analysis artifact (e.g., a visualization) using principles of human-centered design

These are the outcomes most closely related to the Common Core State Standards for Mathematics (CCSSM); the Guidelines for Assessment and Instruction in Statistics Education II (GAISE II), a framework for statistics education that is foundational for current work in developing national data science learning progressions and curriculum; and the Next Generation Science Standards (NGSS).

Design Recommendations

Based on the alignment among the CCSSM, GAISE II, and NGSS, we have the following recommendations for the revised CSTA Computer Science standards in Data and Analysis.

Grade Band	Design Recommendations
All Grades	Explicitly reference statistical problem-solving as a multi-step process in the standards and systematically build students' capacity in all steps. Consider using or adapting the four-step statistical problem-solving process from the GAISE II: (1) formulate statistical investigative questions, (2) collect/consider the data, (3) analyze the data, and (4) interpret the results. Students should read, write, and speak with other students about data and analysis to build fluency with technical terms and how they are used in the context of data and analysis (e.g., "variable" in a dataset vs. a computer program vs. a mathematical equation) and develop their abilities to comprehend, reference, evaluate, and produce informative/explanatory and argumentative text and oral/multimedia presentations about statistical investigations.

Grade Band	Design Recommendations
Elementary	Students should ask and answer questions of small datasets collected in familiar contexts. Questions will generally focus on a single attribute, such as the lengths of objects or the numbers of objects in different categories. Students can collect the data by hand, through counting, categorizing, and measuring objects, or use data that has already been collected. Students should distinguish between quantitative and qualitative data (e.g., words vs. numbers) and know when to use each. Answers can be given in response to structured prompts. Students should begin using simple computational tools to make basic data tables, picture graphs, bar graphs, and line graphs.
Middle	Students should ask and answer questions of large datasets (e.g., hundreds or thousands of cases) that cannot be easily summarized manually and that are collected from a variety of contexts, including science, social science, and humanities. Students should be able to pose a question involving one or two variables in the data, describe how the data were collected, use concepts from math and statistics to generate descriptive summary statistics for the variables of interest (e.g., measures of center and variability, calculating percentages), make informal interpretations, and acknowledge the limitations of their investigation. Students should also use these datasets to reinforce their understanding of concepts they are learning in math that are more readily illustrated with larger datasets. Students should use computational tools to collect and store data (e.g., through online surveys or sensors). Students should begin developing understanding of spreadsheets and how computational tools can assist with transforming, summarizing, and representing/visualizing data (especially using tables, bar charts, line graphs, and scatterplots).
High	Students should ask and answer complex questions involving large datasets collected from a variety of contexts, including science, social science, and humanities. These questions can involve multiple variables/merging of multiple datasets. Students should develop and execute plans for collecting, cleaning, and analyzing their data using computational tools. Students should visualize their data and apply algebraic thinking to data analysis (e.g., fitting linear and non-linear functions). Students should formally report out their findings, either in written reports or in presentations. Students should use spreadsheets and other computational tools to transform, analyze, and visualize data and create simulations. Students should develop test cases to ensure computations and simulations perform as intended. Advanced students may begin using statistical software that requires programming to clean data, conduct descriptive and inferential statistical analysis, and create/modify simulations.

For more information on these design recommendations, go to https://csteachers.org/k12standards/research. See the following sections of *Comparison of CSTA Standards to Other Standards and Frameworks*: Common Core State Standards for Mathematics (pp. 22–28), Guidelines for Assessment and Instruction in Statistics Education II (pp. 29–33), Next Generation Science Standards (pp. 33–38), and Data and Analysis (pp. 65–67).

Data and Analysis

Aligned Learning Progressions

Aligned Data and Analysis Learning Progressions for CCSSM Content Standards[1], GAISE II Framework[2], and NGSS Science and Engineering Practices (SEPs)[3]

The learning progressions in the following pages paraphrase the most closely related concepts and practices for Data and Analysis and are not intended to be a comprehensive catalog of standards, practices, and instructional recommendations for each grade level or grade band.

	Summary of Related CCSSM Content Standards	Summary of Related GAISE II Instructional Recommendations	Summary of NGSS SEPs 4 and 5
K	Count and describe, compare, and categorize objects based on their attributes.	*All recommendations that align to elementary school CCSSM are Level A.* **I. Formulate Statistical Investigative Questions** Understand and experience different types of statistical investigative questions. Pose statistical investigative questions involving familiar contexts and small datasets. **IV. Interpret Results** Use evidence to answer statistical investigative questions and communicate results with teacher guidance.	**SEP 4: Analyzing and interpreting data** Record and share observations. Use observations to describe patterns and relationships to answer questions. **SEP 5: Using mathematics and computational thinking** Decide when to use qualitative vs. quantitative data. Use counting and numbers to identify and describe patterns. Describe, measure, and compare quantitative attributes of objects. Display data using simple graphs.
1	Measure and compare/order objects using arbitrary length units (e.g., a paper clip). Organize and represent data with up to three categories, answering data questions about the number of data points and comparing the number of data points in each category.		
2	Measure and estimate lengths using standard whole units. Create line plots to represent measurement data. Create picture graphs and bar graphs to represent data with up to four categories and answer data questions using information in the graphs (e.g., comparing the quantities in two categories).		

[1] Focusing on the **Measurement and Data** and **Statistics and Probability** domains. There are other related domains within the CCSSM.

[2] The GAISE II framework organizes its instructional recommendations around a four-step statistical problem-solving process.

[3] Focusing on SEP 4: ***Analyzing and interpreting data*** and SEP 5: ***Using mathematics and computational thinking*** because they are most closely related to data and analysis. There are other related practices within the NGSS.

	Summary of Related CCSSM Content Standards	Summary of Related GAISE II Instructional Recommendations	Summary of NGSS SEPs 4 and 5
3	Measure and estimate lengths using half and quarter units. Create line plots to represent measurement data. Create scaled picture graphs and bar graphs to represent data with several categories, and answer "how many more/less" questions using the graphs.	*All recommendations that align to elementary school CCSSM are Level A.* **I. Formulate Statistical Investigative Questions** Understand and experience different types of statistical investigative questions. Pose statistical investigative questions involving familiar contexts and small datasets. **IV. Interpret Results** Use evidence to answer statistical investigative questions and communicate results with teacher guidance.	**SEP 4: Analyzing and interpreting data** Represent data in tables and graphical displays to reveal patterns and relationships. When possible and feasible, use digital tools to analyze and interpret data. **SEP 5: Using mathematics and computational thinking** Decide whether to use quantitative or qualitative data. Organize simple datasets to reveal patterns. Describe, measure, and/or graph quantities related to scientific questions and engineering problems.
4	Create line plots to represent measurement data using fractional units (½, ¼, ⅛). Solve addition and subtraction problems using information in the line plots.		
5	Create line plots to represent measurement data using fractional units (½, ¼, ⅛). Solve problems requiring fraction operations using information in the line plots.		

	Summary of Related CCSSM Content Standards	Summary of Related GAISE II Instructional Recommendations	Summary of NGSS SEPs 4 and 5
6	Recognize statistical questions, which are questions that anticipate variability in the data and account for it in the answers.[4] Summarize datasets by reporting the numbers of observations, describing the attributes in the data and how they were measured, and calculating quantitative measures of center and variability.	**I. Formulate Statistical Investigative Questions** ***Level B:*** Understand that there are different types of statistical investigative questions and they can be used to articulate research topics. Pose statistical investigative questions involving large datasets, including questions about a population, changes over time, and data collected from devices. **II. Collect Data/Consider Data** ***Level A:*** Understand the basic properties of datasets and the variables within them and how data can be collected. Interrogate datasets to understand the variables and how they relate to statistical investigative questions. ***Level B:*** Understand that data are collected with a purpose and can be organized and stored in a variety of structures (e.g., spreadsheets). Recognize the limitations of datasets and data collection methods. Interrogate datasets to determine how the data were collected and understand the variables in the data. Understand how sampling and random assignment are used to answer statistical investigative questions. **III. Analyze the Data** ***Level A:*** Understand and represent data distributions for categorical and quantitative variables. Describe key features of data distributions for quantitative variables, such as measures of center, variability, and shape. Compare distributions and identify associations between variables. ***Level B:*** Visualize data distributions for quantitative variables. Use measures of center, variability, and shape to reason about two distributions and compare them. Explore patterns of association between two variables. **IV. Interpret Results** ***Level B:*** Use statistical evidence to answer statistical investigative questions and communicate results comprehensively, with some teacher guidance. State the limitations of the sample and recognize the uncertainty in generalizing beyond the sample. Compare results for different conditions in an experiment.	**SEP 4: Analyzing and interpreting data** Construct and use graphical displays of large datasets to identify relationships between variables. Use concepts from statistics and probability, including measures of center and variability, to analyze and characterize data. Consider the limitations of datasets and data analysis. Distinguish between causal and correlational relationships. **SEP 5: Using mathematics and computational thinking** Use digital tools (e.g., computers) to analyze very large datasets for patterns and trends. Use mathematical representations and apply mathematical concepts and/or processes (e.g., ratio, rate, percent[5]) to scientific questions and engineering problems.
7	Understand that statistics can be used to learn about a population by examining a sample of the population and generalizing from the sample if the sample is representative. Draw multiple random samples from a dataset to gauge the variability in estimates and predictions from each sample. Use the data from a random sample to draw an inference about a population. Make informal comparative inferences about two populations, such as by assessing the degree of visual overlap between two data distributions and comparing measures of center and variability.		
8	Create scatterplots to investigate associations between two variables, and describe the relationships. When scatterplots suggest a linear relationship, model the relationship by informally fitting a straight line to the points and assess the model fit by judging the closeness of the data points to the line.		

[4] The example provided in the CCSSM ("How old am I?") is not a statistical question because there is a single numerical answer. "How old are the students in my school?" is a statistical question because the answer must contend with the variability in students' ages.

[5] Ratios, rates, and percents are introduced in sixth grade in the CCSSM.

	Summary of Related CCSSM Content Standards	Summary of Related GAISE II Instructional Recommendations	Summary of NGSS SEPs 4 and 5
HS	Represent data with plots on the real number line (dot plots, histograms, and box plots). Create scatterplots with two quantitative variables and describe their relationship. Graph linear and exponential functions. Distinguish between situations that can be modeled using linear and exponential functions. Understand statistics as a process for making inferences about population parameters based on a random sample from that population. Recognize the different purposes of sample surveys, experiments, and observational studies. Use data from sample surveys or randomized experiments to draw inferences. Evaluate reports based on data.	*All recommendations that align to high school CCSSM are Level C.* **I. Formulate Statistical Investigative Questions** Pose complex statistical investigative questions, such as those involving multiple variables or requiring causal inference. Determine how data can be collected and analyzed to answer the questions. **II. Collect Data/Consider Data** Distinguish among different types of statistical investigations, such as surveys, observational studies, and experiments. Understand good practices for sampling, data collection, and data handling. Apply practices when collecting primary data or selecting data for secondary analysis. Understand complex considerations around research design, such as the role of random sampling in surveys, random assignment in experiments, issues around bias and confounding variables, and the generality of findings. **III. Analyze the Data** Use technology to subset, filter, and transform data. Summarize and describe relationships among multiple variables by visualizing the data and calculating summary statistics and confidence intervals. Use simulations to understand and approximate sampling distributions and to investigate associations between categorical variables. **IV. Interpret Results** Use statistical evidence to answer statistical investigative questions and communicate results through formal reports and presentations. Understand how missing data and outliers may influence an analysis. Understand how to interpret margins of error and *p*-values.	**SEP 4: Analyzing and interpreting data** Analyze data using tools, technologies, or models. Analyze data using concepts from statistical probability, such as calculating correlation coefficients and best-fit lines. Consider the limitations of datasets and data analysis. Distinguish between causal and correlational relationships. **SEP 5: Using mathematics and computational thinking** Use algebraic thinking in analysis (e.g., fitting linear and non-linear functions). Create and/or revise computational models or simulations of a phenomenon, device, process, or system. Test mathematical expressions, computer programs, algorithms, or simulations to see if they "make sense" by comparing their outcomes to the real world.