

SECTION SEVEN: Draft High School Data Science and Statistics

7A: Core Data Science and Statistics Focus

The standards listed in this table name the priority instructional content for high school functions (HSF). The right-hand column contains draft focus content that would be core content for all students in a student's first two credits after K-8 mathematics. Specific modeling standards are indicated by a star symbol (★).

HSS.ID – Interpreting Categorical & Quantitative Data

Standard	Standard Statements (Jan 2021 Draft)
HSS.ID.A.1	Represent the distribution of data multiple ways with plots on the real number line.
HSS.ID.A.2	Use statistics appropriate to the shape of the data distribution to compare center and spread of two or more different data sets.
HSS.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).
HSS.ID.A.4	Use the mean and standard deviation of an approximately normally distributed data set to estimate population percentages.
HSS.ID.B.5	Analyze the association between two categorical variables by using two-way tables and comparative bar graphs.
HSS.ID.B.6	Represent data on two quantitative variables on a scatter plot and describe how the variables are related.
HSS.ID.C.7	Interpret the slope (rate of change) and the intercept (constant term) of a linear model in the context of the data.
HSS.ID.C.8	Compute, using technology, and interpret the correlation coefficient of a linear fit.
HSS.ID.C.9	Distinguish between correlation and causation.

HSS.IC – Making Inferences & Justifying Conclusions

Standard	Standard Statements (Jan 2021 Draft)
HSS.IC.A.1	Understand the process of statistical reasoning, formulate questions, collect, analyze, and interpret data to answer statistical investigative questions.
HSS.IC.B.3	Recognize the difference between sample surveys, experiments and observational studies and understand the role of randomization in each.
HSS.IC.B.4	Use data from a sample survey to estimate a population parameter.
HSS.IC.B.5	Use data from a randomized experiment to compare two treatments to decide if differences between parameters are significant based on the statistics.
HSS.IC.B.6	Evaluate reports based on data.

HSS.CP – Conditional Probability & the Rules of Probability

Standard	Standard Statements (Jan 2021 Draft)
HSS.CP.A.1	Describe the possible outcomes for a situation as subsets of a sample space.
HSS.CP.A.5	Recognize and explain the concepts of conditional probability and independence in everyday language and everyday situations. Integrated with Standard(s): HSS.CP.A.4

7B: Remaining Data Science and Statistics Considerations

The concepts listed in this table represent remaining content that is often taught in high school but should only be attended to if students demonstrate proficiency in priority content. The right-hand column contains considerations where this content could be included, integrated, or excluded as well as reference standards for the identified remaining concepts.

Concept	Core Alignment Consideration (January 2021 Draft)
Simulations	Eliminate lessons using simulations to develop a margin of error or decide if differences between parameters are significant. Reference Standard(s): HSS.IC.B.4, HSS.IC.B.5
Independent Events	Limit lessons to conceptual understanding; Eliminate product of probabilities. Reference Standard(s): HSS.CP.A.2, HSS.CP.A.3
Conditional Probability	Limit lessons to conceptual understanding; Eliminate lessons on computation of conditional probabilities. Reference Standard(s): HSS.CP.A.3; HSS.CP.B.6
Addition Rule	Eliminate lessons on applying the addition rule. Reference Standard(s): HSS.CP.B.7
Multiplication Rule	Eliminate lessons on applying the multiplication rule. Reference Standard(s): HSS.CP.B.8
Permutations and Combinations	Limit lessons to conceptual understanding; Eliminate lessons on computation of permutations and combinations. Reference Standard(s): HSS.CP.B.9

7C: High School Data Science and Statistics Crosswalk with Clarifying Guidance

CLUSTER: HSS.ID – Interpreting Categorical & Quantitative Data

STANDARD: HSS.ID.A.1

Original CCSS Text (2010):

Represent data with plots on the real number line (dot plots, histograms, and box plots).

Standards Statement (JAN 2021):

Represent the distribution of data multiple ways with plots on the real number line.

Clarifying Guidance (JAN 2021):

Graph numerical data on a real number line using dot plots, histograms, and box plots. Data are displayed visually to discover patterns and deviations from patterns. Analyze the strengths and weakness inherent in each type of plot by comparing different plots of the same data. Describe and give simple conclusions and interpretations of a graphical representation of data.

STANDARD: HSS.ID.A.2

Original CCSS Text (2010):

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.

Standards Statement (JAN 2021):

Use statistics appropriate to the shape of the data distribution to compare center and spread of two or more different data sets.

Clarifying Guidance (JAN 2021):

Quantitative data can be described in terms of key characteristics: measures of shape, center, and spread. The shape of a data distribution might be described as symmetric, skewed, uniform, or bell shaped, and it might be summarized by a statistic measuring center (such as mean or median) and a statistic measuring spread (such as standard deviation or interquartile range). Students should have the opportunity to gain an understanding of this concept through the use of technology tools.

STANDARD: HSS.ID.A.3

Original CCSS Text (2010):

Interpret differences in shape, center, and spread in the context of the data sets, accounting for possible effects of extreme data points (outliers).

Standards Statement (JAN 2021):

Interpret differences in shape, center, and spread in the context of the data sets, accounting for possible effects of extreme data points (outliers).

Clarifying Guidance (JAN 2021):

Use data from multiple sources to interpret differences in shape, center and spread. Discuss the effect of outliers on measures of center and spread. Students should use spreadsheets, graphing utilities and statistical software to identify outliers and analyze data sets with and without outliers as appropriate.

STANDARD: HSS.ID.A.4

Original CCSS Text (2010):

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.

Standards Statement (JAN 2021):

Use the mean and standard deviation of an approximately normally distributed data set to estimate population percentages.

Clarifying Guidance (JAN 2021):

Data may be displayed using histograms, dot plots, or smooth normal curves. Recognize that there are data sets for which the empirical rule is not appropriate. The use of calculators, spreadsheets, z-score tables, to estimate the area under the curve is not appropriate for the first two years of study.

STANDARD: HSS.ID.B.5

Original CCSS Text (2010):

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.

Standards Statement (JAN 2021):

Analyze the association between two categorical variables by using two-way tables and comparative bar graphs.

Clarifying Guidance (JAN 2021):

Read, interpret and write clear summaries of data displayed in a two-way frequency table. Calculate joint, marginal, and conditional relative frequencies. Make appropriate displays of joint, marginal, and conditional distributions. Describe patterns observed in the data. Recognize the association between two variables by comparing conditional and marginal percentages. Students may use spreadsheets, graphing calculators, and statistical software to create frequency tables and determine associations or trends in the data.

STANDARD: HSS.ID.B.6

Original CCSS Text (2010):

Represent data on two quantitative variables on a scatter plot, and describe how the variables are related.

- | | |
|-------------|--|
| HSS.ID.B.6a | Fit a function to the data; use functions fitted to data to solve problems in the context of the data. Use given functions or choose a function suggested by the context. Emphasize linear, quadratic, and exponential models. |
| HSS.ID.B.6b | Informally assess the fit of a function by plotting and analyzing residuals. |
| HSS.ID.B.6c | Fit a linear function for a scatter plot that suggests a linear association. |

Standards Statement (JAN 2021):

Represent data on two quantitative variables on a scatter plot and describe how the variables are related.

Clarifying Guidance (JAN 2021):

This is a good opportunity for students to collect and graph their own data and use modeling to fit a function to the data; use a function fitted to data to solve problems in the context of the data. (Emphasize linear models.)

Fit a linear function for a scatter plot that suggests a linear association.

Students should use spreadsheets, graphing calculators, and statistical software to analyze the bivariate data.

STANDARD: HSS.ID.C.7

Original CCSS Text (2010):

Interpret the slope (rate of change) and the intercept (constant term) of a linear model in the context of the data.

Standards Statement (JAN 2021):

Interpret the slope (rate of change) and the intercept (constant term) of a linear model in the context of the data.

Clarifying Guidance (JAN 2021):

Students demonstrate interpreting slope in the context of a given situation when examining two variable statistics as being “for each additional known unit increase in an explanatory variable, we expect or predict a known unit increase (or decrease) in the response variable.”

Students demonstrate interpreting intercept in the context of a given situation when examining two variable statistics as being “the predicted known unit of a response variable when the explanatory variable is zero known units.”

Students would use technology to develop an awareness of how outliers might affect the rate of change and the intercept of a given model.

Students should be able to explain when intercepts might be outside the scope of the model.

STANDARD: HSS.ID.C.8

Original CCSS Text (2010):

Compute (using technology) and interpret the correlation coefficient of a linear fit.

Standards Statement (JAN 2021):

Compute, using technology, and interpret the correlation coefficient of a linear fit.

Clarifying Guidance (JAN 2021):

Explain that the correlation coefficient must be between -1 and 1 inclusive and explain what each of these values means. Determine whether the correlation coefficient shows a weak positive, strong positive, weak negative, strong negative, or no linear correlation. Interpret what the correlation coefficient is telling about the data. Students should use spreadsheets, graphing calculators and statistical software to represent data, describe how the variables are related, fit functions to data, perform regressions, and calculate residuals and correlation coefficients.

STANDARD: HSS.ID.C.9

Original CCSS Text (2010):

Distinguish between correlation and causation.

Standards Statement (JAN 2021):

Distinguish between correlation and causation.

Clarifying Guidance (JAN 2021):

Understand and explain the difference between correlation and causation. Understand and explain that a strong correlation does not mean causation. Determine if statements of causation seem reasonable or unreasonable and justify reasoning.

CLUSTER: HSS.IC – Making Inferences & Justifying Conclusions

STANDARD: HSS.IC.A.1

Original CCSS Text (2010):

Understand statistics as a process for making inferences about population parameters based on a random sample from that population.

Standards Statement (JAN 2021):

Understand the process of statistical reasoning, formulate questions, collect, analyze, and interpret data to answer statistical investigative questions.

Clarifying Guidance (JAN 2021):

This is an opportunity for students to create a survey, collect data, and use graphical displays, sample statistics or two way tables to help estimate population parameters which are unknown values. It is important to understand samples used on social media or in the news.

STANDARD: HSS.IC.B.3

Original CCSS Text (2010):

Recognize the purposes of and differences among sample surveys, experiments, and observational studies; explain how randomization relates to each.

Standards Statement (JAN 2021):

Recognize the difference between sample surveys, experiments and observational studies and understand the role of randomization in each.

Clarifying Guidance (JAN 2021):

[no additional clarifying guidance at this time]

STANDARD: HSS.IC.B.4

Original CCSS Text (2010):

Use data from a sample survey to estimate a population mean or proportion; develop a margin of error through the use of simulation models for random sampling.

Standards Statement (JAN 2021):

Use data from a sample survey to estimate a population parameter.

Clarifying Guidance (JAN 2021):

This is an opportunity for students to look at real data, margin of error and discuss what it means to estimate a population parameter.

STANDARD: HSS.IC.B.5

Original CCSS Text (2010):

Use data from a randomized experiment to compare two treatments; use simulations to decide if differences between parameters are significant.

Standards Statement (JAN 2021):

Use data from a randomized experiment to compare two treatments to decide if differences between parameters are significant based on the statistics.

Clarifying Guidance (JAN 2021):

Limit to population proportion, graphical representations, and visual overlap.

STANDARD: HSS.IC.B.6

Original CCSS Text (2010):

Evaluate reports based on data.

Standards Statement (JAN 2021):

Evaluate reports based on data.

Clarifying Guidance (JAN 2021):

[no additional clarifying guidance at this time]

CLUSTER: HSS.CP – Conditional Probability & the Rules of Probability

STANDARD: HSS.CP.A.1

Original CCSS Text (2010):

Describe events as subsets of a sample space (the set of outcomes) using characteristics (or categories) of the outcomes, or as unions, intersections, or complements of other events (“or,” “and,” “not”).

Standards Statement (JAN 2021):

Describe the possible outcomes for a situation as subsets of a sample space.

Clarifying Guidance (JAN 2021):

This provides an opportunity for students to engage with finding the outcomes of situations which include words such as **and**, **or**, **not**, **if**, and **all**, and to grammatical constructions that reflect logical connections.

STANDARD: HSS.CP.A.5

Original CCSS Text (2010):

Recognize and explain the concepts of conditional probability and independence in everyday language and everyday situations.

Standards Statement (JAN 2021):

Recognize and explain the concepts of conditional probability and independence in everyday language and everyday situations.

Clarifying Guidance (JAN 2021):

[no additional clarifying guidance at this time]