Data Concept Development Grades K-5 and Statistics and Probability for Grades 6-8

The following chart shows the Common Core State Standards by grade level based on the concept of Data. It is not designed to show teaching strategies or how standards should be grouped. It is designed to quickly show how the development of a concept progresses across grades K-8.

Kindergarten	1st	2nd
✓ *Classify objects and count the number of objects in each category.* K.MD.3 Classify objects or people into given categories; count the numbers in each category and sort the categories by count. (Note: Limit category counts to be less than or equal to 10.)	✓ *Represent and interpret data.* 1.MD.4 Organize, represent, and interpret data with up to three categories; ask and answer questions about the total number of data points, how many in each category, and how many more or less are in one category than in another.	✓ *Represent and interpret data.* 2.MD.9 Generate measurement data by measuring lengths of several objects to the nearest whole unit, or by making repeated measurements of the same object. Show the measurements by making a line plot, where the horizontal scale is marked off in whole-number units. 2.MD.10 Draw a picture graph and a bar graph (with single-unit scale) to represent a data set with up to four categories. Solve simple put together, take-apart, and compare problems using information presented in a bar graph. (Note: See Glossary, Table 1.)

3rd	4th	5th
✓ *Represent and interpret data.* 3.MD.3 Draw a scaled picture graph and a scaled bar graph to represent a data set with several categories. Solve one- and two-step "how many more" and "how many less" problems using information presented in scaled bar graphs. *For example, draw a bar graph in which each square in the bar graph might represent 5 pets.* 3.MD.4 Generate measurement data by measuring lengths using rulers marked with halves and fourths of an inch. Show the data by making a line plot, where the horizontal scale is marked off in appropriate units—whole numbers, halves, or quarters.	✓ *Represent and interpret data.* 4.MD.4 Make a line plot to display a data set of measurements in fractions of a unit (1/2, 1/4, 1/8). Solve problems involving addition and subtraction of fractions by using information presented in line plots. *For example, from a line plot find and interpret the difference in length between the longest and shortest specimens in an insect collection.*	✓ *Represent and interpret data.* 5.MD.2 Make a line plot to display a data set of measurements in fractions of a unit (1/2, 1/4, 1/8). Use operations on fractions for this grade to solve problems involving information presented in line plots. *For example, given different measurements of liquid in identical beakers, find the amount of liquid each beaker would contain if the total amount in all the beakers were redistributed equally.*

6th	7th	8th
In middle school the concept of *Data* is addressed in the Domain "Statistics and Probability".	In middle school the concept of *Data* is addressed in the Domain "Statistics and Probability".	In middle school the concept of *Data* is addressed in the Domain "Statistics and Probability".

✓ *Develop understanding of statistical variability.*
6.SP.1 Recognize a statistical question as one that anticipates variability in the data related to the question and accounts for it in the answers. *For example, "How old am I?" is not a statistical question, but "How old are the students in my school?" is a statistical question because one anticipates variability in students' ages.*
6.SP.2 Understand that a set of data collected to answer a statistical question has a distribution which can be described by its center, spread, and overall shape.
6.SP.3 Recognize that a measure of center for a numerical data set summarizes all of its values with a single number, while a measure of variation describes how its values vary with a single number.
 	
✓ *Summarize and describe distributions.*
6.SP.4 Display numerical data in plots on a number line, including dot plots, histograms, and box plots.
6.SP.5 Summarize numerical data sets in relation to their context, such as by:
  a. Reporting the number of observations.
  b. Describing the nature of the attribute under investigation, including how it was measured and its units of
	✓ *Use random sampling to draw inferences about a population.*
	7.SP.1 Understand that statistics can be used to gain information about a population by examining a sample of the population; generalizations about a population from a sample are valid only if the sample is representative of that population. Understand that random sampling tends to produce representative samples and support valid inferences.
	7.SP.2 Use data from a random sample to draw inferences about a population with an unknown characteristic of interest. Generate multiple samples (or simulated samples) of the same size to gauge the variation in estimates or predictions. *For example, estimate the mean word length in a book by randomly sampling words from the book; predict the winner of a school election based on randomly sampled survey data. Gauge how far off the estimate or prediction might be.*
	
	✓ *Draw informal comparative inferences about two populations.*
	7.SP.3 Informally assess the degree of visual overlap of two numerical data distributions with similar variabilities, measuring the difference between the centers by expressing it as a multiple of a measure of variability. *For example,*
		✓ *Investigate patterns of association in bivariate data.*
		8.SP.1 Construct and interpret scatter plots for bivariate measurement data to investigate patterns of association between two quantities. Describe patterns such as clustering, outliers, positive or negative association, linear association, and nonlinear association.
		8.SP.2 Know that straight lines are widely used to model relationships between two quantitative variables. For scatter plots that suggest a linear association, informally fit a straight line, and informally assess the model fit by judging the closeness of the data points to the line.
		8.SP.3 Use the equation of a linear model to solve problems in the context of bivariate measurement data, interpreting the slope and intercept. *For example, in a linear model for a biology experiment, interpret a slope of 1.5 cm/hr as meaning that an additional hour of sunlight each day is associated with an additional 1.5 cm in mature plant height.*
		8.SP.4 Understand that patterns of association can also be seen in bivariate categorical data by displaying frequencies and relative frequencies in a two-way table. Construct and interpret a two-way table summarizing data on two categorical variables

measurement.
c. Giving quantitative measures of center (median and/or mean) and variability (interquartile range and/or mean absolute deviation), as well as describing any overall pattern and any striking deviations from the overall pattern with reference to the context in which the data were gathered.
d. Relating the choice of measures of center and variability to the shape of the data distribution and the context in which the data were gathered.

the mean height of players on the basketball team is 10 cm greater than the mean height of players on the soccer team, about twice the variability (mean absolute deviation) on either team; on a dot plot, the separation between the two distributions of heights is noticeable.

7.SP.4 Use measures of center and measures of variability for numerical data from random samples to draw informal comparative inferences about two populations. *For example, decide whether the words in a chapter of a seventh-grade science book are generally longer than the words in a chapter of a fourth-grade science book.*

✓ *Investigate chance processes and develop, use, and evaluate probability models.*

7.SP.5 Understand that the probability of a chance event is a number between 0 and 1 that expresses the likelihood of the event occurring. Larger numbers indicate greater likelihood. A probability near 0 indicates an unlikely event, a probability around 1/2 indicates an event that is neither unlikely nor likely, and a probability near 1 indicates a likely event.

7.SP.6 Approximate the probability of a chance event by collecting data on the chance process that produces it and observing its long-run relative frequency, and predict the approximate relative frequency given the probability. *For example, when rolling a number cube 600 times, predict that a 3 or 6 would be rolled roughly 200 times, but probably not exactly 200 times.*

collected from the same subjects. Use relative frequencies calculated for rows or columns to describe possible association between the two variables. *For example, collect data from students in your class on whether or not they have a curfew on school nights and whether or not they have assigned chores at home. Is there evidence that those who have a curfew also tend to have chores?*

7.SP.7 Develop a probability model and use it to find probabilities of events. Compare probabilities from a model to observed frequencies; if the agreement is not good, explain possible sources of the discrepancy.

a. Develop a uniform probability model by assigning equal probability to all outcomes, and use the model to determine probabilities of events. *For example, if a student is selected at random from a class, find the probability that Jane will be selected and the probability that a girl will be selected.*

b. Develop a probability model (which may not be uniform) by observing frequencies in data generated from a chance process. *For example, find the approximate probability that a spinning penny will land heads up or that a tossed paper cup will land open-end down. Do the outcomes for the spinning penny appear to be equally likely based on the observed frequencies?*

7.SP.8 Find probabilities of compound events using organized lists, tables, tree diagrams, and simulation.

a. Understand that, just as with simple events, the probability of a compound event is the fraction of outcomes in the sample space for which the compound event occurs.

b. Represent sample spaces for compound events using methods such as organized lists, tables and tree diagrams. For an event described in everyday language (e.g., "rolling double sixes"), identify the

	outcomes in the sample space which compose the event. c. Design and use a simulation to generate frequencies for compound events. *For example, use random digits as a simulation tool to approximate the answer to the question: If 40% of donors have type A blood, what is the probability that it will take at least 4 donors to find one with type A blood?*	