Mathematics for College Statistics B.E.S.T. Instructional Guide for Mathematics

The B.E.S.T. Instructional Guide for Mathematics (B1G-M) is intended to assist educators with planning for student learning and instruction aligned to Florida's Benchmarks for Excellent Student Thinking (B.E.S.T.) Standards. This guide is designed to aid high-quality instruction through the identification of components that support the learning and teaching of the B.E.S.T. Mathematics Standards and Benchmarks. The B1G-M includes an analysis of information related to the B.E.S.T. Standards for Mathematics within this specific mathematics course, the instructional emphasis and aligned resources. This document is posted on the B.E.S.T. Standards for Mathematics webpage of the Florida Department of Education's website and will continue to undergo edits as needed.

Structural Framework and Intentional Design of the B.E.S.T. Standards for Mathematics

Florida's B.E.S.T. Standards for Mathematics were built on the following.

- The coding scheme for the standards and benchmarks was changed to be consistent with other content areas. The new coding scheme is structured as follows: Content.GradeLevel.Strand.Standard.Benchmark.
- Strands were streamlined to be more consistent throughout.
- The standards and benchmarks were written to be clear and concise to ensure that they are easily understood by all stakeholders.
- The benchmarks were written to allow teachers to meet students' individual skills, knowledge and ability.
- The benchmarks were written to allow students the flexibility to solve problems using a method or strategy that is accurate, generalizable and efficient depending on the content (i.e., the numbers, expressions or equations).
- The benchmarks were written to allow for student discovery (i.e., exploring) of strategies rather than the teaching, naming and assessing of each strategy individually.
- The benchmarks were written to support multiple pathways for success in career and college for students.
- The benchmarks should not be taught in isolation but should be combined purposefully.
- The benchmarks may be addressed at multiple points throughout the year, with the intention of gaining mastery by the end of the year.
- Appropriate progression of content within and across strands was developed for each grade level and across grade levels.
- There is an intentional balance of conceptual understanding and procedural fluency with the application of accurate real-world context intertwined within mathematical concepts for relevance.
- The use of other content areas, like science and the arts, within real-world problems should be accurate, relevant, authentic and reflect grade-level appropriateness.

Components of the B.E.S.T. Instructional Guide for Mathematics

The following table is an example of the layout for each benchmark and includes the defining attributes for each component. It is important to note that instruction should not be limited to the possible connecting benchmarks, related terms, strategies or examples provided. To do so would strip the intention of an educator meeting students' individual skills, knowledge and abilities.

Benchmark
focal point for instruction within lesson or task

This section includes the benchmark as identified in the B.E.S.T. Standards for Mathematics. The benchmark, also referred to as the Benchmark of Focus, is the focal point for student learning and instruction. The benchmark, and its related example(s) and clarification(s), can also be found in the course description. The 9-12 benchmarks may be included in multiple courses; select the example(s) or clarification(s) as appropriate for the identified course.

Connecting Benchmarks/Horizontal Alignment
in other standards within the grade level or course

This section includes a list of connecting benchmarks that relate horizontally to the Benchmark of Focus. Horizontal alignment is the intentional progression of content within a grade level or course linking skills within and across strands. Connecting benchmarks are benchmarks that either make a mathematical connection or include prerequisite skills. The information included in this section is not a comprehensive list, and educators are encouraged to find other connecting benchmarks. Additionally, this list will not include benchmarks from the same standard since benchmarks within the same standard already have an inherent connection.

Terms from the K-12 Glossary

This section includes terms from Appendix C: K-12 Glossary, found within the B.E.S.T. Standards for Mathematics document, which are relevant to the identified Benchmark of Focus. The terms included in this section should not be viewed as a comprehensive vocabulary list, but instead should be considered during instruction or act as a reference for educators.

Vertical Alignment
across grade levels or courses

This section includes a list of related benchmarks that connect vertically to the Benchmark of Focus. Vertical alignment is the intentional progression of content from one year to the next, spanning across multiple grade levels. Benchmarks listed in this section make mathematical connections from prior grade levels or courses in future grade levels or courses within and across strands. If the Benchmark of Focus is a new concept or skill, it may not have any previous benchmarks listed. Likewise, if the Benchmark of Focus is a mathematical skill or concept that is finalized in learning and does not have any direct connection to future grade levels or courses, it may not have any future benchmarks listed. The information included in this section is not a comprehensive list, and educators are encouraged to find other benchmarks within a vertical progression.

Purpose and Instructional Strategies

This section includes further narrative for instruction of the benchmark and vertical alignment. Additionally, this section may also include the following:

- explanations and details for the benchmark;
- vocabulary not provided within Appendix C;
- possible instructional strategies and teaching methods; and
- strategies to embed potentially related Mathematical Thinking and Reasoning Standards (MTRs).

Common Misconceptions or Errors

This section will include common student misconceptions or errors and may include strategies to address the identified misconception or error. Recognition of these misconceptions and errors enables educators to identify them in the classroom and make efforts to correct the misconception or error. This corrective effort in the classroom can also be a form of formative assessment within instruction.

Instructional Tasks

demonstrate the depth of the benchmark and the connection to the related benchmarks

This section will include example instructional tasks, which may be open-ended and are intended to demonstrate the depth of the benchmark. Some instructional tasks include integration of the Mathematical Thinking and Reasoning Standards (MTRs) and related benchmark(s). Enrichment tasks may be included to make connections to benchmarks in later grade levels or courses. Tasks may require extended time, additional materials and collaboration.

Instructional Items

demonstrate the focus of the benchmark

This section will include example instructional items which may be used as evidence to demonstrate the students' understanding of the benchmark. Items may highlight one or more parts of the benchmark.

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

Mathematical Thinking and Reasoning Standards

MTRs: Because Math Matters

Florida students are expected to engage with mathematics through the Mathematical Thinking and Reasoning Standards (MTRs) by utilizing their language as a self-monitoring tool in the classroom, promoting deeper learning and understanding of mathematics. The MTRs are standards which should be used as a lens when planning for student learning and instruction of the B.E.S.T. Standards for Mathematics.

Structural Framework and Intentional Design of the Mathematical Thinking and Reasoning Standards

The Mathematical Thinking and Reasoning Standards (MTRs) are built on the following.

- The MTRs have the same coding scheme as the standards and benchmarks; however, they are written at the standard level because there are no benchmarks.
- In order to fulfill Florida's unique coding scheme, the 5th place (benchmark) will always be a "1" for the MTRs.
- The B.E.S.T. Standards for Mathematics should be taught through the lens of the MTRs.
- At least one of the MTRs should be authentically and appropriately embedded throughout every lesson based on the expectation of the benchmark(s).
- The bulleted language of the MTRs were written for students to use as self-monitoring tools during daily instruction.
- The clarifications of the MTRs were written for teachers to use as a guide to inform their instructional practices.
- The MTRs ensure that students stay engaged, persevere in tasks, share their thinking, balance conceptual understanding and procedures, assess their solutions, make connections to previous learning and extended knowledge, and apply mathematical concepts to real-world applications.
- The MTRs should not stand alone as a separate focus for instruction, but should be combined purposefully.
- The MTRs will be addressed at multiple points throughout the year, with the intention of gaining mastery of mathematical skills by the end of the year and building upon these skills as they continue in their K-12 education.

MA.K12.MTR.1.1 Actively participate in effortful learning both individually and collectively.

Mathematicians who participate in effortful learning both individually and with others:

- Analyze the problem in a way that makes sense given the task.
- Ask questions that will help with solving the task.
- Build perseverance by modifying methods as needed while solving a challenging task.
- Stay engaged and maintain a positive mindset when working to solve tasks.
- Help and support each other when attempting a new method or approach.

Clarifications:
Teachers who encourage students to participate actively in effortful learning both individually and with others:

- Cultivate a community of growth mindset learners.
- Foster perseverance in students by choosing tasks that are challenging.
- Develop students' ability to analyze and problem solve.
- Recognize students' effort when solving challenging problems.

MA.K12.MTR.2.1 Demonstrate understanding by representing problems in multiple ways.

Mathematicians who demonstrate understanding by representing problems in multiple ways:

- Build understanding through modeling and using manipulatives.
- Represent solutions to problems in multiple ways using objects, drawings, tables, graphs and equations.
- Progress from modeling problems with objects and drawings to using algorithms and equations.
- Express connections between concepts and representations.
- Choose a representation based on the given context or purpose.

Clarifications:
Teachers who encourage students to demonstrate understanding by representing problems in multiple ways:

- Help students make connections between concepts and representations.
- Provide opportunities for students to use manipulatives when investigating concepts.
- Guide students from concrete to pictorial to abstract representations as understanding progresses.
- Show students that various representations can have different purposes and can be useful in different situations.

MA.K12.MTR.3.1 Complete tasks with mathematical fluency.

Mathematicians who complete tasks with mathematical fluency:

- Select efficient and appropriate methods for solving problems within the given context.
- Maintain flexibility and accuracy while performing procedures and mental calculations.
- Complete tasks accurately and with confidence.
- Adapt procedures to apply them to a new context.
- Use feedback to improve efficiency when performing calculations.

Clarifications:
Teachers who encourage students to complete tasks with mathematical fluency:

- Provide students with the flexibility to solve problems by selecting a procedure that allows them to solve efficiently and accurately.
- Offer multiple opportunities for students to practice efficient and generalizable methods.
- Provide opportunities for students to reflect on the method they used and determine if a more efficient method could have been used.

MA.K12.MTR.4.1 Engage in discussions that reflect on the mathematical thinking of self and others.

Mathematicians who engage in discussions that reflect on the mathematical thinking of self and others:

- Communicate mathematical ideas, vocabulary and methods effectively.
- Analyze the mathematical thinking of others.
- Compare the efficiency of a method to those expressed by others.
- Recognize errors and suggest how to correctly solve the task.
- Justify results by explaining methods and processes.
- Construct possible arguments based on evidence.

Clarifications:
Teachers who encourage students to engage in discussions that reflect on the mathematical thinking of self and others:

- Establish a culture in which students ask questions of the teacher and their peers, and error is an opportunity for learning.
- Create opportunities for students to discuss their thinking with peers.
- Select, sequence and present student work to advance and deepen understanding of correct and increasingly efficient methods.
- Develop students' ability to justify methods and compare their responses to the responses of their peers.

MA.K12.MTR.5.1 Use patterns and structure to help understand and connect mathematical concepts.

Mathematicians who use patterns and structure to help understand and connect mathematical concepts:

- Focus on relevant details within a problem.
- Create plans and procedures to logically order events, steps or ideas to solve problems.
- Decompose a complex problem into manageable parts.
- Relate previously learned concepts to new concepts.
- Look for similarities among problems.
- Connect solutions of problems to more complicated large-scale situations.

Clarifications:
Teachers who encourage students to use patterns and structure to help understand and connect mathematical concepts:

- Help students recognize the patterns in the world around them and connect these patterns to mathematical concepts.
- Support students to develop generalizations based on the similarities found among problems.
- Provide opportunities for students to create plans and procedures to solve problems.
- Develop students' ability to construct relationships between their current understanding and more sophisticated ways of thinking.

MA.K12.MTR.6.1 Assess the reasonableness of solutions.

Mathematicians who assess the reasonableness of solutions:

- Estimate to discover possible solutions.
- Use benchmark quantities to determine if a solution makes sense.
- Check calculations when solving problems.
- Verify possible solutions by explaining the methods used.
- Evaluate results based on the given context.

Clarifications:
Teachers who encourage students to assess the reasonableness of solutions:

- Have students estimate or predict solutions prior to solving.
- Prompt students to continually ask, "Does this solution make sense? How do you know?"
- Reinforce that students check their work as they progress within and after a task.
- Strengthen students' ability to verify solutions through justifications.

MA.K12.MTR.7.1 Apply mathematics to real-world contexts.

Mathematicians who apply mathematics to real-world contexts:

- Connect mathematical concepts to everyday experiences.
- Use models and methods to understand, represent and solve problems.
- Perform investigations to gather data or determine if a method is appropriate.
- Redesign models and methods to improve accuracy or efficiency.

Clarifications:
Teachers who encourage students to apply mathematics to real-world contexts:

- Provide opportunities for students to create models, both concrete and abstract, and perform investigations.
- Challenge students to question the accuracy of their models and methods.
- Support students as they validate conclusions by comparing them to the given situation.
- Indicate how various concepts can be applied to other disciplines.

Examples of Teacher and Student Moves for the MTRs

Below are examples that demonstrate the embedding of the MTRs within the mathematics classroom. The provided teacher and student moves are examples of how some MTRs could be incorporated into student learning and instruction keeping in mind the benchmark(s) that are the focal point of the lesson or task. The information included in this table is not a comprehensive list, and educators are encouraged to incorporate other teacher and student moves that support the MTRs.

MTR	Student Moves	Teacher Moves
MA.K12.MTR.1.1 *Actively participate in effortful learning both individually and collectively.*	• Students engage in the task through individual analysis, student-to-teacher interaction and student-to-student interaction. • Students ask task-appropriate questions to self, the teacher and to other students. *(MTR.4.1)* • Students have a positive productive struggle exhibiting growth mindset, even when making a mistake. • Students stay engaged in the task to a purposeful conclusion while modifying methods, when necessary, in solving a problem through self-analysis and perseverance.	• Teacher provides flexible options (i.e., differentiated, challenging tasks that allow students to actively pursue a solution both individually and in groups) so that all students have the opportunity to access and engage with instruction, as well as demonstrate their learning. • Teacher creates a physical environment that supports a growth mindset and will ensure positive student engagement and collaboration. • Teacher provides constructive, encouraging feedback to students that recognizes their efforts and the value of analysis and revision. • Teacher provides appropriate time for student processing, productive struggle and reflection. • Teacher uses data and questions to focus students on their thinking; help students determine their sources of struggle and to build understanding. • Teacher encourages students to ask appropriate questions of other students and of the teacher including questions that examine accuracy. *(MTR.4.1)*

MTR	Student Moves	Teacher Moves
MA.K12.MTR.2.1 *Demonstrate understanding by representing problems in multiple ways.*	• Students represent problems concretely using objects, models and manipulatives. • Students represent problems pictorially using drawings, models, tables and graphs. • Students represent problems abstractly using numerical or algebraic expressions and equations. • Students make connections and select among different representations and methods for the same problem, as appropriate to different situations or context. *(MTR.3.1)*	• Teacher provides students with objects, models, manipulatives, appropriate technology and real-world situations. *(MTR.7.1)* • Teacher encourages students to use drawings, models, tables, expressions, equations and graphs to represent problems and solutions. • Teacher questions students about making connections between different representations and methods and challenges students to choose one that is most appropriate to the context. *(MTR.3.1)* • Teacher encourages students to explain their different representations and methods to each other. *(MTR.4.1)* • Teacher provides opportunities for students to choose appropriate methods and to use mathematical technology.
MA.K12.MTR.3.1 *Complete tasks with mathematical fluency.*	• Students complete tasks with flexibility, efficiency and accuracy. • Students use feedback from peers and teachers to reflect on and revise methods used. • Students build confidence through practice in a variety of contexts and problems. *(MTR.1.1)*	• Teacher provides tasks and opportunities to explore and share different methods to solve problems. *(MTR.1.1)* • Teacher provides opportunities for students to choose methods and reflect (i.e., through error analysis, revision, summarizing methods or writing) on the efficiency and accuracy of the method(s) chosen. • Teacher asks questions and gives feedback to focus student thinking to build efficiency of accurate methods. • Teacher offers multiple opportunities to practice generalizable methods.

MTR	Student Moves	Teacher Moves
MA.K12.MTR.4.1 *Engage in discussions that reflect on the mathematical thinking of self and others.*	• Students use content specific language to communicate and justify mathematical ideas and chosen methods. • Students use discussions and reflections to recognize errors and revise their thinking. • Students use discussions to analyze the mathematical thinking of others. • Students identify errors within their own work and then determine possible reasons and potential corrections. • When working in small groups, students recognize errors of their peers and offers suggestions.	• Teacher provides students with opportunities (through open-ended tasks, questions and class structure) to make sense of their thinking. *(MTR.1.1)* • Teacher uses precise mathematical language, both written and abstract, and encourages students to revise their language through discussion. • Teacher creates opportunities for students to discuss and reflect on their choice of methods, their errors and revisions and their justifications. • Teachers select, sequence and present student work to elicit discussion about different methods and representations. *(MTR.2.1, MTR.3.1)*

MTR	Student Moves	Teacher Moves
MA.K12.MTR.5.1 *Use patterns and structure to help understand and connect mathematical concepts.*	• Students identify relevant details in a problem in order to create plans and decompose problems into manageable parts. • Students find similarities and common structures, or patterns, between problems in order to solve related and more complex problems using prior knowledge.	• Teacher asks questions to help students construct relationships between familiar and unfamiliar problems and to transfer this relationship to solve other problems. *(MTR.1.1)* • Teacher provides students opportunities to connect prior and current understanding to new concepts. • Teacher provides opportunities for students to discuss and develop generalizations about a mathematical concept. *(MTR.3.1, MTR.4.1)* • Teacher allows students to develop an appropriate sequence of steps in solving problems. • Teacher provides opportunities for students to reflect during problem solving to make connections to problems in other contexts, noticing structure and making improvements to their process.
MA.K12.MTR.6.1 *Assess the reasonableness of solutions.*	• Students estimate a solution, including using benchmark quantities in place of the original numbers in a problem. • Students monitor calculations, procedures and intermediate results during the process of solving problems. • Students verify and check if solutions are viable, or reasonable, within the context or situation. *(MTR.7.1)* • Students reflect on the accuracy of their estimations and their solutions.	• Teacher provides opportunities for students to estimate or predict solutions prior to solving. • Teacher encourages students to compare results to estimations and revise if necessary for future situations. *(MTR.5.1)* • Teacher prompts students to self-monitor by continually asking, "Does this solution or intermediate result make sense? How do you know?" • Teacher encourages students to provide explanations and justifications for results to self and others. *(MTR.4.1)*

MTR	Student Moves	Teacher Moves
MA.K12.MTR.7.1 *Apply mathematics to real-world contexts.*	• Students connect mathematical concepts to everyday experiences. • Students use mathematical models and methods to understand, represent and solve real-world problems. • Students investigate, research and gather data to determine if a mathematical model is appropriate for a given situation from the world around them. • Students re-design models and methods to improve accuracy or efficiency.	• Teacher provides real-world context to help students build understanding of abstract mathematical ideas. • Teacher encourages students to assess the validity and accuracy of mathematical models and situations in real-world context, and to revise those models if necessary. • Teacher provides opportunities for students to investigate, research and gather data to determine if a mathematical model is appropriate for a given situation from the world around them. • Teacher provides opportunities for students to apply concepts to other content areas.

Mathematics for College Statistics Areas of Emphasis

In Mathematics for College Statistics, instructional time will emphasize four areas:

(1) analyzing and applying linear and exponential functions within the context of statistics;
(2) extending understanding of probability using data and various representations, including two-way tables and Venn Diagrams;
(3) representing and interpreting univariate and bivariate categorical and numerical data; and
(4) determining the appropriateness of different types of statistical studies.

The purpose of the areas of emphasis is not to guide specific units of learning and instruction, but rather provide insight on major mathematical topics that will be covered within this mathematics course. In addition to its purpose, the areas of emphasis are built on the following.

- Supports the intentional horizontal progression within the strands and across the strands in this grade level or course.
- Student learning and instruction should not focus on the stated areas of emphasis as individual units.
- Areas of emphasis are addressed within standards and benchmarks throughout the course so that students are making connections throughout the school year.
- Some benchmarks can be organized within more than one area.
- Supports the communication of the major mathematical topics to all stakeholders.
- Benchmarks within the areas of emphasis should not be taught within the order in which they appear. To do so would strip the progression of mathematical ideas and miss the opportunity to enhance horizontal progressions within the grade level or course.

The table on the next page shows how the benchmarks within this mathematics course are embedded within the areas of emphasis.

		Analyzing and Applying Linear and Exponential Functions	Extending Understanding of Probability Using Data and Various Representations	Representing and Interpreting Various Data Types	Determining the Appropriateness of Different Types of Statistical Studies
Algebraic Reasoning	MA.912.AR.1.1	x	x		
	MA.912.AR.1.2	x	x		
	MA.912.AR.2.5	x			
	MA.912.AR.5.7	x			
Functions	MA.912.F.1.2	x	x		
	MA.912.F.1.8	x		x	
Financial Literacy	MA.912.FL.1.1	x	x		
	MA.912.FL.1.3		x	x	
Data Analysis & Probability	MA.912.DP.1.1	x	x	x	
	MA.912.DP.1.2	x	x	x	
	MA.912.DP.1.3	x	x	x	
	MA.912.DP.2.1			x	
	MA.912.DP.2.4	x		x	
	MA.912.DP.2.5	x		x	
	MA.912.DP.2.6	x		x	
	MA.912.DP.2.7	x		x	
	MA.912.DP.2.9	x		x	
	MA.912.DP.3.1		x	x	
	MA.912.DP.3.2		x	x	
	MA.912.DP.3.5		x	x	
	MA.912.DP.4.1		x		
	MA.912.DP.4.2		x		
	MA.912.DP.4.3		x		
	MA.912.DP.4.4		x		
	MA.912.DP.4.5		x		

		Analyzing and Applying Linear and Exponential Functions	Extending Understanding of Probability Using Data and Various Representations	Representing and Interpreting Various Data Types	Determining the Appropriateness of Different Types of Statistical Studies
	MA.912.DP.4.6		x		
	MA.912.DP.4.7		x		
	MA.912.DP.4.8		x		
	MA.912.DP.4.9		x		
	MA.912.DP.4.10		x		
	MA.912.DP.5.1			x	x
	MA.912.DP.5.2				x
	MA.912.DP.5.3				x
	MA.912.DP.5.4			x	x
	MA.912.DP.5.5	x	x	x	x
	MA.912.DP.5.6				x
	MA.912.DP.5.7				x
	MA.912.DP.5.11	x	x	x	x
Logic & Discrete Theory	MA.912.LT.5.4		x		
	MA.912.LT.5.5		x		

Algebraic Reasoning

MA.912.AR.1 *Interpret and rewrite algebraic expressions and equations in equivalent forms.*

MA.912.AR.1.1

Benchmark

MA.912.AR.1.1	Identify and interpret parts of an equation or expression that represent a quantity in terms of a mathematical or real-world context, including viewing one or more of its parts as a single entity.

Algebra I Example: Derrick is using the formula $P = 1000(1+.1)^t$ to make a prediction about the camel population in Australia. He identifies the growth factor as $(1+.1)$, or 1.1, and states that the camel population will grow at an annual rate of 10% per year.

Example: The expression 1.15^t can be rewritten as $\left(1.15^{\frac{1}{12}}\right)^{12t}$ which is approximately equivalent to 1.012^{12t}. This latter expression reveals the approximate equivalent monthly interest rate of 1.2% if the annual rate is 15%.

Benchmark Clarifications:

Clarification 1: Parts of an expression include factors, terms, constants, coefficients and variables.

Clarification 2: Within the Mathematics for Data and Financial Literacy course, problem types focus on money and business.

Connecting Benchmarks/Horizontal Alignment

- MA.912.AR.2.5
- MA.912.AR.5.7
- MA.912.DP.2.4, MA.912.DP.2.9
- MA.912.DP.4.7, MA.912.DP.4.8, MA.912.DP.4.10

Terms from the K-12 Glossary

- Coefficient
- Expression
- Equation

Vertical Alignment

Previous Benchmarks

- MA.912.NSO.1.2
- MA.912.AR.2.2, MA.912.AR.2.6
- MA.912.AR.3.1, MA.912.AR.3.6, MA.912.AR.3.7, MA.912.AR.3.8
- MA.912.AR.4.1
- MA.912.AR.5.3, MA.912.AR.5.6
- MA.912.FL.3.2

Next Benchmarks

Purpose and Instructional Strategies

In Algebra I, students generated and interpreted equivalent linear, absolute value, quadratic and exponential expressions, and equations. In Mathematics for College Statistics, students continue to apply this concept to all types of equations in order to assist in interpreting data and situations.

- Instruction includes identifying and interpreting parts of equations for both statistics equations and probability equations and how they relate to the context that they represent.
- The expectation in this course is that students are using the form $y = a + bx$ where a is the y-intercept and b is the slope when writing the equation of the line of best fit for data sets that show a linear trend. This is different from the typical slope-intercept form of $y = mx + b$ that is already familiar to students.
- Problem types are limited to real-world context.

Common Misconceptions or Errors

- Students may not be able to identify parts of an expression and equation or interpret those parts within context. Ensure these are embedded throughout instruction and discussions.
  - For example, building in questions to identify these parts and discussing their connection to the context in which they represent in a routine way will help students to make these connections.
- Students may confuse the slope, b, and the y-intercept, a, when using $y = a + bx$ due to students not being familiar with this format.

Instructional Tasks

Instructional Task 1 (MTR.7.1)

The contingency table below shows how a sample of high school students at Carver High School get to school on a typical morning. The data is also broken down by sex.

	Rides the Bus	Parent Drives	Student Drives	Student Walks or Bikes	Total
Boys	10	14	5	3	32
Girls	4	15	7	1	27
Total	14	29	12	4	59

Part A. Let A be the event that the student drives to school. Let B be the probability that the student rides the bus to school. Using the addition rule for probability ($P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$), replace A and B with the events given above. Interpret each part of three parts of the addition rule for probability. Calculate each of the three parts of the addition rule for probability. What is $P(\text{the student drives to school or the student rides the bus to school})$?

Part B. Let A be the event that a parent drives the student to school. Let B be the probability that the student is female. Using the addition rule for probability ($P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$), replace A and B with the events given above. Interpret each part of three parts of the addition rule for probability. Calculate each of the three parts of the addition rule for probability. What is $P(\text{a parent drives the student or the student is female})$?

Instructional Items

Instructional Item 1

A simple random sample of 50 high school students is used to collect data on GPA and SAT mathematics scores. Suppose the scatter plot of data is linear and produces a regression equation of $Predicted\ SAT = 550.8 + 56.3\ (GPA)$.

Part A. Interpret the y-intercept of the equation. Does it make sense in this context?

Part B. Interpret the slope of the equation.

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

MA.912.AR.1.2

Benchmark

MA.912.AR.1.2 Rearrange equations or formulas to isolate a quantity of interest.

Algebra I Example: The Ideal Gas Law $PV = nRT$ can be rearranged as $T = \frac{PV}{nR}$ to isolate temperature as the quantity of interest.

Example: Given the Compound Interest formula $= P\ (1 + \frac{r}{n})^{nt}$, solve for P.

Mathematics for Data and Financial Literacy Honors Example: Given the Compound Interest formula $= P\ (1 + \frac{r}{n})^{nt}$, solve for t.

Benchmark Clarifications:

Clarification 1: Instruction includes using formulas for temperature, perimeter, area and volume; using equations for linear (standard, slope-intercept and point-slope forms) and quadratic (standard, factored and vertex forms) functions.

Clarification 2: Within the Mathematics for Data and Financial Literacy course, problem types focus on money and business.

Connecting Benchmarks/Horizontal Alignment

- MA.912.AR.2.5
- MA.912.AR.5.7
- MA.912.DP.2.4, MA.912.DP.2.9
- MA.912.DP.4.7, MA.912.DP.4.8, MA.912.DP.4.10

Terms from the K-12 Glossary

- Coefficient
- Expression
- Equation
- Linear expression (or linear equation)
- Linear function

Vertical Alignment

Previous Benchmarks

- MA.912.NSO.1.2
- MA.912.AR.2.2, MA.912.AR.2.6
- MA.912.AR.3.1, MA.912.AR.3.6, MA.912.AR.3.7, MA.912.AR.3.8
- MA.912.AR.4.1
- MA.912.AR.5.3, MA.912.AR.5.6
- MA.912.FL.3.2

Next Benchmarks

Purpose and Instructional Strategies

In Algebra I, students rearranged equations and formulas to isolate a specific variable. In Mathematics for College Statistics, students continue to apply this concept to all types of equations in order to assist in interpreting data and situations.

- Instruction includes identifying parts of equations for both statistics equations and probability equations and how they relate to the context that they represent.
- Instruction focuses on rearranging linear and exponential functions as they relate to quantitative bivariate data.
- Problem types are limited to real-world context.
- Instruction focuses on rearranging formulas used for probability in order to solve for unknowns.
    - For example, the formula $P(A \cap B) = P(A) \cdot P(B)$ is used to check if events are independent. If you know events A and B are independent and they have the probability of event B, they can rearrange the equation to solve for $P(A)$.

Common Misconceptions or Errors

- Students may mistake the variable $P(A)$ as two separate variables because it is using two letters when involving probability.
- Students may be confused by the hat above the y-variable in a linear function used to estimate the best fit of the data.
- Students may not be able to confidently work with inverse operations especially when fractions are involved.

Instructional Tasks

Instructional Task 1 (MTR.4.1)

Using the addition rule for probability $P(A\ or\ B) = P(A) + P(B) - P(A\ and\ B)$

Part A. What formula could be used to find the $P(A\ and\ B)$?

Part B. Let $A = Female$ and $B = Comedy$, explain why the formula for $P(A\ and\ B)$ from above makes sense.

	Comedy	Action	Total
Male	8	12	20
Female	8	4	12
Total	16	16	32

Part C. How would the formula be different if the events A and B were mutually exclusive? Compare your answer with a partner and come to a consensus.

Instructional Items

Instructional Item 1

A simple random sample of 50 high school students is used to collect data on GPA and SAT mathematics scores. Suppose the scatter plot of data is linear and produces a regression equation of $Predicted\ SAT = 550.8 + 56.3\ (GPA)$.

Part A. Rewrite the equation by solving for GPA.

Part B. Suppose someone uses this equation to predict a score of 660 on the mathematics portion of the SAT. Use your equation above to find their GPA.

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

MA.912.AR.2 *Write, solve and graph linear equations, functions and inequalities in one and two variables.*

MA.912.AR.2.5

Benchmark

MA.912.AR.2.5	Solve and graph mathematical and real-world problems that are modeled with linear functions. Interpret key features and determine constraints in terms of the context.

Algebra I Example: Lizzy's mother uses the function $C(p) = 450 + 7.75p$, where $C(p)$ represents the total cost of a rental space and p is the number of people attending, to help budget Lizzy's 16th birthday party. Lizzy's mom wants to spend no more than \$850 for the party. Graph the function in terms of the context.

Benchmark Clarifications:

Clarification 1: Key features are limited to domain, range, intercepts and rate of change.

Clarification 2: Instruction includes the use of standard form, slope-intercept form and point-slope form.

Clarification 3: Instruction includes representing the domain, range and constraints with inequality notation, interval notation or set-builder notation.

Clarification 4: Within the Algebra I course, notations for domain and range are limited to inequality and set-builder.

Clarification 5: Within the Mathematics for Data and Financial Literacy course, problem types focus on money and business.

Connecting Benchmarks/Horizontal Alignment	Terms from the K-12 Glossary
- MA.912.AR.5.7 - MA.912.DP.2.4 - MA.912.F.1.2, MA.912.F.1.8	- Coordinate Plane - Domain - Function notation - Range - Rate of change - Slope x-intercept y-intercept

Vertical Alignment

Previous Benchmarks	Next Benchmarks
- MA.8.AR.3.4, MA.8.AR.3.5 - MA.912.AR.2.4	

Purpose and Instructional Strategies

In Algebra I, students solved mathematical and real-world problems that were modeled with linear functions using various linear models including standard form, point-slope form and slope intercept form. These linear situations were also presented in various ways such as in tables or written descriptions. Additionally, students interpreted key features such as rate of change, intercepts, domain and range while also determining constraints. In Mathematics for College Statistics, students interpret key features and determine possible constraints when analyzing real-world data in order to examine bivariate data.

- This benchmark is an opportunity for students to review what they had done previously in Algebra I. Students should be able to fine tune their ability to interpret key features such as rate of change, intercepts, domain and range along with constraints in context of actual real world data. This will allow students to see how they can apply these techniques with actual real world scenarios prior to exploring this in more depth when they begin working with the Data Analysis and Probability benchmarks.
- Instruction includes representing domain, range and constraints using words, inequality notation, set-builder notation and interval notation.
    - **Words**
      If the domain is all real numbers, it can be written as "all real numbers" or "any value of x, such that x is a real number."
    - **Inequality Notation**
      If the domain is all values of x greater than 2, it can be represented as $x > 2$.
    - **Set-Builder Notation**
      If the domain is all values of x less than or equal to zero, it can be represented as $\{x|x \leq 0\}$ and is read as "all values of x such that x is less than or equal to zero."
    - **Interval Notation**
      If the domain is all values from 0 to infinity, including 0, it can be represented as $[0, \infty)$.
- Depending on a student's pathway, they may not have worked with interval notation (as that was not an expectation in Algebra I) before this course. Instruction includes making connections between inequality notation and interval notation.

- For example, if the range of a function is $-10 < y < 24$, it can be represented in interval notation as $(-10, 24)$. This is commonly referred to as an open interval because the interval does not contain the end values.
  - For example, if the domain of a function is $0 \leq x \leq 11.5$, it can be represented in interval notation as $[0, 11.5]$. This is commonly referred to as a closed interval because the interval contains both end values.
  - For example, if the domain of a function is $0 \leq x < 50$, it can be represented in interval notation as $[0, 50)$. This is commonly referred to as a half-open, or half-closed, interval because the interval contains only one of the end values.
  - For example, if the range of a function is all real numbers, is can be represented in interval notation as $(-\infty, \infty)$. This is commonly referred to as an infinite interval because at least one of end values is infinity (positive or negative).
- The expectation in this course is that students are using the form $y = a + bx$ where a is the y-intercept and b is the slope when writing the equation of the line of best fit for data sets that show a linear trend. This is different from the typical slope-intercept form of $y = mx + b$ that is already familiar to students.
- Instruction includes focusing on graphing the line of best fit for a set of bivariate data, how the line fits into the data, the interpretation of the slope and y-intercept, and the use of the linear regression equation to make predictions.
- Instruction includes the use of statistics programs or graphing calculators in order to allow students to graph varying situations in order to focus on interpreting and solving scenarios that can be modeled by a linear function.
- Instruction includes the use of real-world data or scenarios to make connections about with how interpreting key features and using constraints can be helpful when analyzing a real world scenario.
- Instruction includes the understanding that a real-world context can be represented by a linear two-variable equation even though it only has meaning for discrete values.
  - For example, if a gym membership cost $10.00 plus $6.00 for each class, this can be represented as $y = 10 + 6c$. When represented on the coordinate plane, the relationship is graphed using the points (0,10), (1,16), (2,22) and so on.
- Instruction directs students to graph or interpret a representation of a context that necessitates a constraint. Discuss the meaning of multiple points on the line and announce their meanings in the associated context *(MTR.4.1)*. Allow students to discover that some points do not make sense in context and therefore should not be included in a formal solution *(MTR.6.1)*. Ask students to determine which parts of the line create sensible solutions and guide them to make constraints to represent these sections.
- Problem types are limited to real-world context.

Common Misconceptions or Errors

- Students may find it difficult to make inferences when analyzing bivariate data that may fit a linear model if they do not have the basic concepts of modeling with linear functions, interpreting key features of linear functions and being able to determine the constraints in terms of the context that is given.
- Students may express initial confusion with the meaning of $f(x)$ for functions written in function notation. To help address this, consider writing the same function in both forms simultaneously (*MTR 2.1*).

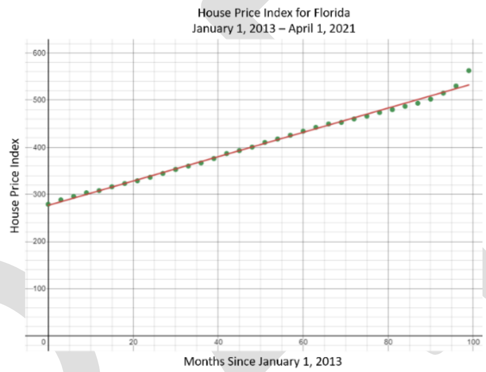
- For example, the function $\frac{2}{3}x + 6$ can be written as $f(x) = \frac{2}{3}x + 6$ and $y = \frac{2}{3}x + 6$, to show that both $f(x)$ and y represent the same outputs of the function.

- Students may assign their constraints to the incorrect variable.
- Students may reverse the values for the slope and y-intercept.
- Students may miss the need for compound inequalities in their constraints. Students may not include zero as part of the domain or range.
  - For example, if a constraint for the domain is between 0 and 10, a student may forget to include 0 in some contexts, since they may assume that one cannot have negative people, for instance.

Instructional Tasks

Instructional Task 1 (MTR.6.1, MTR.7.1)

Shown is the linear regression of the data of the House Price Index for Florida from January 1, 2013 to April 1, 2021. The house price index measures the average price changes in sales on the same properties. The data that was used for this regression was pulled from Economic Research Federal Reserve Bank of St. Louis. (Data pulled from https://fred.stlouisfed.org/series/FLSTHPI)



The function that represents this data set is $H(m) = 276.75 + 2.59m$, where $H(m)$ represents the House Price Index and m represents months since January 1, 2013.

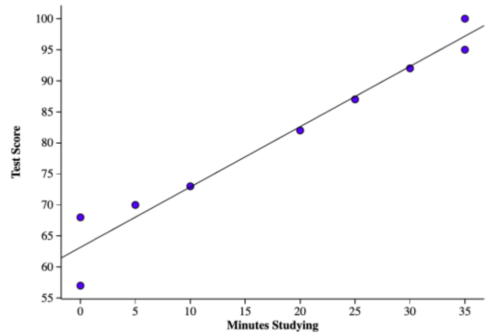
Part A. Interpret the slope and y-intercept within the context of this data set.

Part B. Can you predict what the House Price Index would have been for March 1, 2016? What do you believe the House Price Index would be for December 1, 2021? Does this seem reasonable? Explain your reasoning.

Part C. What are the constraints for this situation? Can we extrapolate outside these constraints? Why or why not?

Instructional Task 2 (MTR.6.1, MTR.7.1)

A group of students test scores and the number of minutes they spent studying for that test have been recorded and a scatter plot and least squares regression line equation is shown below.



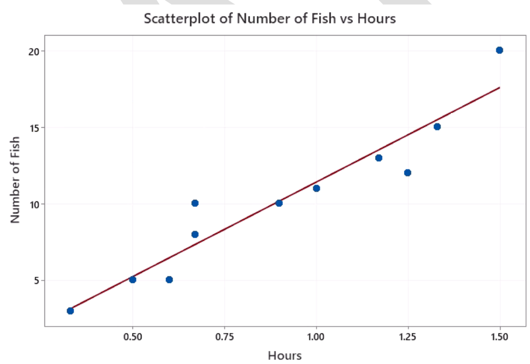
The function that represents this data set is $T(m) = 63.17 + 0.97m$, where $T(m)$ represents the test score and m represents minutes studying.

Part A. Interpret the slope and y-intercept within the context of this data set.

Part B. Predict the test score a student would get if they spent 15 minutes studying. What test score would you predict for someone who spends 190 minutes studying? Does this seem reasonable? Explain.

Instructional Task 3 (MTR.7.1, MTR.6.1)

Over the summer Timothy would go fishing each weekend. He would track how long he fished and the number of fish that he caught. The data and the line of best fit are displayed in the graph below.



The function that represents the regression equation is $f(t) = -0.98 + 12.37t$, where $f(t)$ represents the number of fish Timothy caught and t represents the number of hours he was fishing.

Part A. Interpret the y-intercept in the context of the scenario. Does this make sense? If Timothy were to fish for 0 hours, how many fish would he really expect to catch?

Part B. Interpret the slope in the context of the scenario.

Part C. Suppose Timothy is planning to go fishing this weekend for 45 minutes (.75 hours). Use the regression equation to predict how many fish he will catch.

Instructional Items

Instructional Item 1

The sales of pet food and treats from 2011 to 2021 can be modeled by the following function, $P(x) = 16.36 + 2.58x$, where $P(x)$ represents the sales in billions U.S. dollars and x represents the year since 2011.

Part A. Graph this relationship using a graphing utility.

Part B. Identify the features of this graph and what they mean in context of this data set.

Instructional Item 2

An artist has been tracking their sales from 2015 to 2020 and found that her sales could be modeled by the function $R(x) = 2403 + 3.25x$, where $R(x)$ represent the revenue from sales in U.S. dollars and x represents the year since 2015.

Part A. Graph this relationship using a graphing utility.

Part B. Identify and interpret the following:

- Domain
- Range
- y-intercept
- Slope
- x-intercept

Part C. Are there any values in part B that do not make sense given the context of this situation?

Instructional Item 3

Suppose that Sun Valley High School has a reading club for students in grades 9-12. At the end of the school year, the club sponsor collects data on each member's age and the number of books he or she has read as a member of the reading club. The bivariate data shows a linear trend, and produces a line of best fit with the equation $b(x) = -15.01 + 14.22x$, where $b(x)$ represents the number of books the member has read while in the club and x is the member's age.

Part A. Use technology to graph the regression equation.

Part B. Identify and interpret the features of this graph in context.

Part C. What are reasonable constraints for the domain of data?

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

MA.912.AR.5 *Write, solve and graph exponential and logarithmic equations and functions in one and two variables.*

MA.912.AR.5.7

Benchmark

MA.912.AR.5.7	Solve and graph mathematical and real-world problems that are modeled with exponential functions. Interpret key features and determine constraints in terms of the context.

Example: The graph of the function $f(t) = e^{5t+2}$ can be transformed into the straight line $y = 5t + 2$ by taking the natural logarithm of the function's outputs.

Benchmark Clarifications:

Clarification 1: Key features are limited to domain; range; intercepts; intervals where the function is increasing, decreasing, positive or negative; constant percent rate of change; end behavior and asymptotes.

Clarification 2: Instruction includes representing the domain, range and constraints with inequality notation, interval notation or set-builder notation.

Clarification 3: Instruction includes understanding that when the logarithm of the dependent variable is taken and graphed, the exponential function will be transformed into a linear function.

Clarification 4: Within the Mathematics for Data and Financial Literacy course, problem types focus on money and business.

Connecting Benchmarks/Horizontal Alignment

- MA.912.AR.2.5
- MA.912.F.1.2, MA.912.F.1.8
- MA.912.DP.2.9

Terms from the K-12 Glossary

- Domain
- Exponential Function
- Intercept
- Range
- Rate of Change
- Set-Builder Notation
- x-intercept
- y-intercept

Vertical Alignment

Previous Benchmarks

- MA.912.AR.5.3, MA.912.AR.5.4, MA.912.AR.5.6

Next Benchmarks

Purpose and Instructional Strategies

In Algebra I, students identified and described exponential functions in terms of growth rates or decay rates. Students also wrote exponential functions that modeled relationships characterized by having a constant percent of change per unit interval. In Mathematics for College Statistics, students expand on these ideas by identifying key features within a real-world context using data.

- Instruction includes the use of technology, such as a statistical program or graphing calculators.
- Instruction includes the use of real-world data to allow students to identify and interpret key features within a real-world context. Key features include, but are not limited to, domain, range, intercepts, rate of change, asymptotes and end behavior.
- Instruction includes representing domain, range and constraints using words, inequality notation, set-builder notation and interval notation.
    - **Words**
      If the domain is all real numbers, it can be written as "all real numbers" or "any value of x, such that x is a real number."
    - **Inequality Notation**
      If the domain is all values of x greater than 2, it can be represented as $x > 2$.
    - **Set-Builder Notation**
      If the domain is all values of x less than or equal to zero, it can be represented as $\{x|x \leq 0\}$ and is read as "all values of x such that x is less than or equal to zero."
    - **Interval Notation**
      If the domain is all values from 0 to infinity, including 0, it can be represented as $[0, \infty)$.
- Depending on a student's pathway, they may not have worked with interval notation (as that was not an expectation in Algebra I) before this course. Instruction includes making connections between inequality notation and interval notation.
    - For example, if the range of a function is $-10 < y < 24$, it can be represented in interval notation as $(-10, 24)$. This is commonly referred to as an open interval because the interval does not contain the end values.
    - For example, if the domain of a function is $0 \leq x \leq 11.5$, it can be represented in interval notation as $[0, 11.5]$. This is commonly referred to as a closed interval because the interval contains both end values.
    - For example, if the domain of a function is $0 \leq x < 50$, it can be represented in interval notation as $[0, 50)$. This is commonly referred to as a half-open, or half-closed, interval because the interval contains only one of the end values.
    - For example, if the range of a function is all real numbers, is can be represented in interval notation as $(-\infty, \infty)$. This is commonly referred to as an infinite interval because at least one of end values is infinity (positive or negative).
- Instruction includes comparing and contrasting linear and exponential functions. Students should be able to identify how the rates of changes from these two functions differ from each other.
- When doing statistical analysis on bivariate data, it is sometimes helpful to re-express the data. One situation is when we plot a scatter plot and note that the data is forming a curve, either by the scatter plot itself or the residuals. When this happens, it may be useful to re-express the data using logarithms in order to have the data form a linear

relationship. A linear relationship is easier to analyze when looking at key features in context. Students will need to explore this relationship in order to make the best decision when analyzing data as far as whether there is a need to straighten the data or if the data is straight enough.

- Problem types are limited to real-world context.

Common Misconceptions or Errors

- When describing intervals where functions are increasing, decreasing, positive or negative, students may represent their interval using the incorrect variable. In these cases, ask reflective questions to help students examine the meaning of the domain and range in the problem.
- Students may be unfamiliar with using interval notation. They may miss the need for compound inequalities in their intervals. In these cases, refer to the graph of the function to help them discover areas in their interval that would not make sense in context.
- Students may confuse the rate of change of an exponential function with that of a linear function. Students may not realize that a linear rate of change is a constant rate whereas an exponential rate of change is a constant ratio. It may be helpful for students to see both at the same time and exploring the differences between these two functions as it relates to rate of change.
- Depending on the students' course pathway in grade 9-12, students may not have had experience with changing exponential functions to a linear representation using logarithms.

Instructional Tasks

Instructional Task 1 (MTR.7.1)

Photographers will examine the relationship between aperture and shutter speed to decide what they need for the shot that they are going for. For aperture, the higher the value the smaller the lens opening. This controls the amount of light that is used to expose the film. For shutter speed, the larger the value in the denominator the faster it is. The table below provides data comparing shutter speed to aperture.

Shutter Speed	1/3200	1/1600	1/800	1/400	1/200	1/100	1/50	1/25	1/13	1/6
Aperture	1.4	2	2.8	4	5.6	8	11	16	22	32

Part A. Create a scatter plot of the data. What do you notice? Does a linear fit seem appropriate for this data?

Part B. Take the log of the aperture values and then graph the data of shutter speed versus $log(aperture)$. What do you notice? Does a linear fit seem appropriate for this data?

Part C. Take the log of the shutter speed values and then graph the data of $log(shutter\ speed)$ versus aperture. What do you notice? Does a linear fit seem appropriate for this data?

Part D. Graph the data of $log(shutter\ speed)$ versus $log(aperture)$. What do you notice? Does a linear fit seem appropriate for this data?

Part E. Which of the four graphs that were created do you think is best described as linear? Explain your reasoning.

Instructional Items

Instructional Item 1

Shang-Chi and the Legend of the Ten Rings was released on September 3, 2021 in U.S. Theaters. The graph below shows how much money the movie grossed, in dollars, daily for 29 days. (https://www.boxofficemojo.com/release/rl3490022913/)

The regression equation is $g(x) = 26{,}776{,}000(0.85572)^x$ where $g(x)$ represents the amount the movie grossed, in dollars, and x represents the day since September 3rd.

Part A. What does the rate of change represent in this function? What does the y-intercept represent in context of this situation?

Part B. What is the end behavior of this function? Does it make sense in context of this situation? Why or why not?

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

Functions

MA.912.F.1 *Understand, compare and analyze properties of functions.*

MA.912.F.1.2

Benchmark

MA.912.F.1.2 Given a function represented in function notation, evaluate the function for an input in its domain. For a real-world context, interpret the output.

Algebra I Example: The function $f(x) = \frac{x}{7} - 8$ models Alicia's position in miles relative to a water stand x minutes into a marathon. Evaluate and interpret for a quarter of an hour into the race.

Benchmark Clarifications:

Clarification 1: Problems include simple functions in two-variables, such as $f(x, y) = 3x - 2y$.

Clarification 2: Within the Algebra I course, functions are limited to one-variable such as $f(x) = 3x$.

Connecting Benchmarks/Horizontal Alignment

- MA.912.DP.2.4, MA.912.DP.2.9

Terms from the K-12 Glossary

- Domain
- Function notation
- Range

Vertical Alignment

Previous Benchmarks

Next Benchmarks

- MA.912.AR.2
- MA.912.AR.3.8
- MA.912.AR.4.3
- MA.912.AR.5.6, MA.912.AR.5.8
- MA.912.DP.2.4

Purpose and Instructional Strategies

In Algebra I, students worked with x-y notation and function notation throughout instruction of linear, quadratic, exponential and absolute value functions. In Mathematics for College Statistics, students continue to use this concept of evaluating a function for an input in its domain when making predictions and extrapolating using different types of regression models when examining real-world data.

- Instruction includes showing students the practicality that function notation represents to mathematicians. In several contexts, multiple functions can exist that we want to consider simultaneously. If each of these functions is written in x-y notation, it can lead to confusion in discussions.
  - For example, the equations $y = 4 - 2x$ and $y = 7 + 3x$. Representing these equations in function notation allows mathematicians to distinguish them from each other more easily (i.e., $f(x) = 4 - 2x$ and $g(x) = 7 + 3x$).
- Function notation also allows for the use of different symbols for the variables, which can add meaning to the function.
  - For example, $h(t) = -16t^2 + 49t + 4$ could be used to represent the height, h, of a ball in feet over time, t, in seconds.
- Function notation allows mathematicians to express the output and input of a function simultaneously.
  - For example, $h(3) = 7$ would represent a ball 7 feet in the air after 3 seconds of elapsed time. This is equivalent to the ordered pair $(3,7)$ but with the added benefit of knowing which function it is associated with,
- Instruction includes how function notation makes clear which variables depend on each other.

Common Misconceptions or Errors

- Students may read function notation, $f(x)$, as $f \cdot x$. Continue the discussion of the meaning of function notation so that students become comfortable with its concept.

Instructional Tasks

Instructional Task 1 (MTR.7.1)

A linear regression was created from data obtained from U.S. Energy Information Administration (https://www.eia.gov/dnav/pet/pet_pri_gnd_dcus_nus_a.htm). The data represents the weekly retail price of regular grade gasoline beginning on January 4, 2021, to September 13, 2021. Florida's weekly regular grade gasoline price can be represented by the function $F(x) = 2.41 + 0.019x$, where x is the number of weeks since January 4, 2021, and $F(x)$ is the average cost of regular gasoline in Florida. California's weekly regular grade gasoline price can be represented by the function $C(x) = 3.268 + 0.032x$, where x is the number of weeks since January 4, 2021, and $C(x)$ is the average cost of regular gasoline in California.

Part A. When you examine the functions for the average weekly price for regular gasoline in Florida and in California, what do you notice? What do you wonder?

Part B. What is $C(0)$ and what does that mean in context of this scenario?

Part C. What do you think the average cost of regular gasoline was on August 23, 2021, (33 weeks) for Florida? California?

Part D. The actual average cost of regular gasoline on August 23, 2021, in Florida was \$2.943 and in California it was \$4.232. How did your estimate compare with the actual averages? Why do you think there is a difference? Explain your reasoning.

Part E. Why do you think the average cost of gasoline is higher in California than in Florida? Explain your reasoning.

Instructional Items

Instructional Item 1

The function, $F(x) = 2.41 + 0.019x$, represents the average weekly cost of regular gasoline in Florida, where x is the number of weeks since January 4, 2021, and $F(x)$ is the average cost of regular gasoline in Florida. What would $F(23) = 2.847$ represent?

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

MA.912.F.1.8

Benchmark

MA.912.F.1.8	Determine whether a linear, quadratic or exponential function best models a given real-world situation.

Benchmark Clarifications:

Clarification 1: Instruction includes recognizing that linear functions model situations in which a quantity changes by a constant amount per unit interval; that quadratic functions model situations in which a quantity increases to a maximum, then begins to decrease or a quantity decreases to a minimum, then begins to increase; and that exponential functions model situations in which a quantity grows or decays by a constant percent per unit interval.

Clarification 2: Within this benchmark, the expectation is to identify the type of function from a written description or table.

Connecting Benchmarks/Horizontal Alignment

- MA.912.AR.1.1
- MA.912.AR.2.5
- MA.912.AR.5.7
- MA.912.DP.2.4, MA.912.DP.2.9

Terms from the K-12 Glossary

- Exponent
- Exponential function
- Linear function
- Quadratic function
- Rate of change

Vertical Alignment

Previous Benchmarks

- MA.8.AR.3.2, MA.8.AR.3.3
- MA.8.AR.4.1
- MA.8.F.1.2

Next Benchmarks

Purpose and Instructional Strategies

In grade 8, students worked with linear functions in graph, table, and equation form to determine if they modeled a linear function as well as identified the slope and y-intercept. In Algebra I, students expanded on their knowledge of linear functions to also determine if the function was quadratic or exponential. In Mathematics for College Statistics, students continue to use these concepts to determine the best model to use for regression analysis.

- Instruction includes recognizing that linear functions model situations in which a quantity changes by a constant amount per unit interval.
- Instruction includes recognizing that quadratic functions model situations in which a quantity increases to a maximum, then begins to decrease or a quantity decreases to a minimum, then begins to increase.
- Instruction includes recognizing that exponential functions model situations in which a quantity grows or decays by a constant percent per unit interval.
- Instruction includes representing these functions with hand-drawn graphs and graphs from technology.
- When the independent variable changes by a constant amount, linear functions have constant first differences, quadratic functions have constant second differences, and exponential functions have a constant ratio.
- Student will interpret the x-intercept, y-intercept, and rate of growth or decay of an exponential function given in a real-world context.
- Students should be able to identify linear, quadratic or exponential functions when they are presented in a written description.
    - For example, a car that depreciates by 16% each year would be decreasing exponentially because it would be taking the value and multiplying by a constant value less than 1. You rent a car and after the initial fees you are charged $30 per day, this would be linear because each day the total expense is increasing by the same amount.

Common Misconceptions or Errors

- Students may identify any equation that has x^2 in it as a quadratic.
    - For example, $y = 3x^2 + 2x^4$.
- Students may attempt to determine first differences, second differences, and if the function is changing at a constant rate without taking into consideration if the independent variable is changing by a constant amount.
- Students may try to find the differences in the y-variable of a table of an exponential function by subtracting.
- Students may not pay attention to the x-variables and miss the pattern if the table shows a break in the pattern for the x-variables.
- Students may make simple mistakes when multiplying by a percentage and forget to change the percent to a decimal or do it incorrectly.
- Students may focus on the percent of decrease rather than the percentage remaining in exponential decay problems.
- Students may not account for the previous value in problems with exponential growth, or decay, and simply multiply by the percent of increase, or decrease.
    - For example, if something is growing by 20% per day, students would multiply the previous value by 120% or 1.2.

Instructional Tasks

Instructional Task 1 (MTR.5.1)

Jessica and Kayla are training for an athletic competition and in order to build up endurance they will be jumping rope. They each started with two minutes on the first day and then a certain number of additional minutes each day after for four days. Kayla recorded the number of minutes she spent jumping rope in the table below. Jessica decided that every day she would do twice as many minutes jumping rope as she did the day before.

	Day 1	Day 2	Day 3	Day 4
Number of Minutes Kayla spent jumping rope	2	5	8	11

Part A. Did the number of minutes Jessica spent jumping rope change at a constant rate per day? Show your work and justify your answer.

Part B. Did the number of minutes Kayla spent jumping rope change at a constant rate per day? Show your work and justify your answer.

Part C. Whose rate of minutes spent jumping rope is increasing more rapidly – Jessica's or Kayla's? Justify your answer.

Part D. What type of function would best model the number of minutes spent jumping rope by Kayla: linear, quadratic or exponential? Justify your answer.

Part E. What type of function would best model the number of minutes spent jumping rope by Jessica: linear, quadratic or exponential? Justify your answer.

Instructional Items

Instructional Item 1

State whether each relationship can be modeled by a linear, quadratic or exponential model justify your choice.

a. The relationship between the height of a basketball being thrown into a basket and the time in seconds.
b. The relationship between the distance traveled and the cost of an Uber trip if the driver charges $6 to pick up and $0.50 per mile.
c. The relationship between amount of radioactive substance and time when the radioactive substance doubles each day.

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

Financial Literacy

MA.912.FL.1 *Build mathematical foundations for financial literacy.*

MA.912.FL.1.1

Benchmark

MA.912.FL.1.1 Extend previous knowledge of operations of fractions, percentages and decimals to solve real-world problems involving money and business.

Benchmark Clarifications:
Clarification 1: Problems include discounts, markups, simple interest, tax, tips, fees, percent increase, percent decrease and percent error.

Connecting Benchmarks/Horizontal Alignment

- MA.912.AR.1.1
- MA.912.DP.2.9
- MA.912.DP.3.5
- MA.912.DP.4.3

Terms from the K-12 Glossary

- Dividend
- Divisor
- Percent of change
- Percent error
- Rate of change
- Simple interest
- Unit rates

Vertical Alignment

Previous Benchmarks

- MA.6.NSO.2.3
- MA.7.AR.3.1
- MA.912.FL.3.1

Next Benchmarks

Purpose and Instructional Strategies

In middle grades, students continued to use fractions, decimals, and percentages in a real-world context that included operations of adding, subtracting, multiplying and dividing. Students also converted between forms of fractions, decimals and percentages. If students took Algebra II, they had experience with comparing simple and compound interests and using Annual Percentage Rates (APR) to solve real world problems. In Mathematics for College Statistics, this benchmark is an extension of MA.7.AR.3.1 where students work with real world problems involving money and business that include such things as discounts, markup, simple interest, tax, tips, fees, percent increase, percent decrease and percent error.

- Instruction includes the use of software capable of statistical analysis. This includes but is not limited to spreadsheets or graphing calculators. It is not an expectation that student calculate weighted averages by hand.
- Instruction includes the use of real-world data.
- Instruction involves real-world applications relating to money and business.

Common Misconceptions or Errors

- Students may confuse how to perform operations with fractions and when a common denominator is required.
- Students who convert fractions to decimals may get answers wrong from rounding too much.
- Students often get confused when percentages are very small and written as a decimal (ex. 0.07%).

Instructional Tasks

Instructional Task 1 (MTR.7.1)

Alexis has just graduated college as a registered dietitian and got a job making $58,127 a year. She will have to start making payments on her student loans next month. She owes a total of $35,200 in student loans. Her loan agreement has an annual interest rate of 5.8% with a 96-month term. Calculate her monthly payment and the total long term cost. To do this you will need to use the equation: $\mathrm{P} = \frac{\mathrm{Cr}(1+\mathrm{r})^{\mathrm{N}}}{(1+\mathrm{r})^{\mathrm{N}}-1}$ where $P =$ monthly payment, $C =$ loan amount, $N =$ the number of months and $r =$ monthly interest rate (a 7.5% annual rate would be converted to a decimal and divided by 12; $(\frac{.075}{12} = 0.00625)$.

Part A. Calculate Alexis's monthly payment.
Part B. Calculate the total long-term cost of her loan.
Part C. What percentage of Alexis's net pay is going towards paying her student loan?

Instructional Items

Instructional Item 1

Jake's gross pay for this month is $7,855 and his deductions make up 14.8% of his monthly salary. He keeps $1/3$ of his net pay in savings. Out of the remaining money, he spends 40% on food and 20% on rent. How much does Jake spend on food and rent this month?

Instructional Item 2

Jessica buys a dress that is discounted by 15% on the price tag. A sign on the rack stated, "Additional discount of 10%." This is the same as a discount of what percentage of the original price of the dress?

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

MA.912.FL.1.3

Benchmark

MA.912.FL.1.3 Solve real-world problems involving weighted averages using spreadsheets and other technology.

Example: Kiko wants to buy a new refrigerator and decides on the following rating system: capacity 50%, water filter life 30% and capability with technology 20%. One refrigerator gets 8 (out of 10) for capacity, 6 for water filter life and 7 for capability with technology. Another refrigerator gets 9 for capacity, 4 for water filter life and 6 for capability with technology. Which refrigerator is best based on the rating system?

Connecting Benchmarks/Horizontal Alignment

- MA.912.DP.2.1

Terms from the K-12 Glossary

Vertical Alignment

Previous Benchmarks

- MA.6.DP.1.2
- MA.7.AR.3.1
- MA.7.DP.1.1
- MA.912.DP.1.4
- MA.912.GR.3.1

Next Benchmarks

Purpose and Instructional Strategies

In middle grades, students continued the use of the mean when interpreting data, but also began to determine when the use of the mean would be appropriate based on the shape of the data set. Students also compared the means when looking at two different data sets. This continued in Algebra I where students found and interpreted the mean of the data set within context. In Mathematics for College Statistics, students now explore weighted averages and how this is applied within a real-world context.

- Instruction includes the use of software capable of statistical analysis. This includes but is not limited to spreadsheets or graphing calculators. It is not an expectation that student calculate weighted averages by hand.
- Instruction includes the use of real-world data.
- Instruction includes familiarizing students with mathematical symbols that are used in statistical formulas. The formula for weighted average is $\bar{X} = \frac{\sum_{i=1}^{n}(x_i \cdot w_i)}{\sum_{i=1}^{n} w_i}$ where $\bar{X}$ is the weighted average, x is the value and w is the weight. Additionally, Σ means “the sum of multiple terms.”

Common Misconceptions or Errors

- Students may find the average of the data set neglecting to use the weights that are given. Students need to have an understanding that one cannot assume that all data points carry equal weights. There are times when the data vary in weight.
    - For example, when calculating one's grades, the grades may not all be of equal weight. Especially if the teacher uses categories.
- Students may not understand that Σ means "the sum of multiple terms."

Instructional Tasks

Instructional Task 1

Lana is a hobbyist and is wanting to buy some 2 inch by 4 inch by 4 foot pinewood in order to build candle holders to sell through an online store. The following are what she bought from various lumber stores.

Quantity	Price per Piece
3	$8.79
5	$3.08
7	$3.76
2	$4.54

Part A. Lana is wanting to figure out how much to charge for her candle holder so that the cost of the wood is covered in the price. What was the average price of the 2 inch by 4 inch by 4 foot pinewood?

Part B. How can she use this average in order to determine how much she would want to charge customers for the candle holders that she made?

Instructional Items

Instructional Item 1

You are interested in calculating your grade in a math class. The teacher uses the following grading system:

Tests: 60%
Quizzes: 20%
Homework: 10%
Classwork: 10%

You currently have a test average of 85, a quiz average of 90, a homework average of 100, and a classwork average of 100. What is your current grade in the class?

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

Data Analysis & Probability

MA.912.DP.1 *Summarize, represent and interpret categorical and numerical data with one and two variables.*

MA.912.DP.1.1

Benchmark

MA.912.DP.1.1	Given a set of data, select an appropriate method to represent the data, depending on whether it is numerical or categorical data and on whether it is univariate or bivariate.

Benchmark Clarifications:
Clarification 1: Instruction includes discussions regarding the strengths and weaknesses of each data display.
Clarification 2: Numerical univariate includes histograms, stem-and-leaf plots, box plots and line plots; numerical bivariate includes scatter plots and line graphs; categorical univariate includes bar charts, circle graphs, line plots, frequency tables and relative frequency tables; and categorical bivariate includes segmented bar charts, joint frequency tables and joint relative frequency tables.
Clarification 3: Instruction includes the use of appropriate units and labels and, where appropriate, using technology to create data displays.

Connecting Benchmarks/Horizontal Alignment

- MA.912.DP.2
- MA.912.DP.3

Terms from the K-12 Glossary

- Bar graph
- Bivariate data
- Box plot
- Categorical data
- Central angle
- Circle graph
- Data
- Frequency table
- Histogram
- Joint frequency
- Joint relative frequency
- Line graph
- Line plot
- Scatter plot
- Stem-and-leaf plot

Vertical Alignment

Previous Benchmarks

- MA.7.DP.1
- MA.8.DP.1

Next Benchmarks

Purpose and Instructional Strategies

In Algebra I, students interpreted the components of data displays for numerical and categorical data, both univariate and bivariate. In Mathematics for College Statistics, students continue with this exploration of being able to differentiate between univariate and bivariate data. Along with differentiating between numerical (quantitative) and categorical (qualitative) data.

- Instruction includes the use of terms such as quantitative data for numerical data and qualitative data for categorical data. Later statistics courses may use these terms interchangeably and students should have knowledge of these terms in order to avoid confusion.
- Instruction includes the use of spreadsheets and graphing calculators with statistical software.
- Instruction includes discussions regarding the strengths and weaknesses of each data display.
- The intention of this benchmark is to be taught in tandem with MA.912.DP.1.2. This is to reinforce students understanding of numerical/categorical (quantitative/qualitative) data, univariate/bivariate data sets and their displays.
    - The data displays for numerical (quantitative) univariate data includes histograms (grouped and ungrouped), relative frequency histograms (grouped and ungrouped), cumulative frequency histogram (grouped and ungrouped), stem-and-leaf plots, box plots and line plots.
    - The data displays for numerical (quantitative) bivariate data includes scatter plots and line graphs.
    - The data displays for categorical (qualitative) univariate data includes bar charts, circle graphs, frequency tables and relative frequency tables.
    - The data displays for categorical (qualitative) bivariate data includes side-by-side bar graphs, segmented bar charts, joint frequency tables and joint relative frequency tables.
- Joint frequency tables can also be referred to as two-way tables or contingency tables.
- This benchmark reinforces the importance of the use of appropriate units and labels.

Common Misconceptions or Errors

- Students may confuse bar graphs and histograms.
- Student may put spaces between their bars in a histogram or not have spaces between the bars in a bar graph.
- Students may forget to include stem values where there is a hole in the data, especially with stem-and-leaf plots.
    - For example, if the data included the numbers 52, 54, 61, 63, 67, 68, 83 and 85 they may list them stems as 5, 6 and 8 when they should be including 7 even though there are no leaves in that stem.
- Students may forget to label the axes on their graphical displays.
- Students may misread a stem-and-leaf plot when the key requires numbers that are not simply two-digit integers.
- Student may forget to include a key with a stem-and-leaf plot.
- Students may create graphical displays that are misleading by using a break in the y-axis.
- Students may see variables that are using numbers and assume the data is numerical (quantitative) when the numbers are used as a way to group things into categories. For example: zip codes and student ID numbers.

Instructional Tasks

Instructional Task 1 (MTR.2.1, MTR.4.1)

The data below was taken from a random sample of high school seniors.

Sex	Height (in)	Weight (lbs)	Eye Color	Hair Color	Grade on Last Math Test
Male	72	218	Brown	Brown	87
Female	69	178	Blue	Brown	95
Male	69	165	Green	Blonde	76
Male	71	174	Brown	Blonde	98
Female	65	155	Blue	Blonde	78
Male	73	180	Brown	Brown	100
Male	67	158	Brown	Blonde	90
Female	60	125	Blue	Blonde	65
Male	68	160	Brown	Brown	92
Female	66	162	Blue	Brown	95
Male	67	165	Blue	Brown	80
Male	70	170	Brown	Blonde	92
Male	72	187	Green	Brown	85
Female	64	135	Brown	Brown	75
Female	65	140	Blue	Blonde	98
Female	71	140	Brown	Blonde	100
Male	75	215	Brown	Brown	95
Female	62	168	Green	Red	87

Part A. Create a univariate categorical display. Compare your display with a partner.
Part B. Create a univariate numerical display. Compare your display with a partner.
Part C. Create a bivariate categorical display. Compare your display with a partner.
Part D. Create a bivariate numerical display. Compare your display with a partner.

Instructional Task 2 (MTR.3.1)

For each of the following determine if the data would be numerical (quantitative) or categorical (qualitative), univariate or bivariate, and create an appropriate display to represent the data.

Part A. A 1-week study of a 5-mile road in Wauchula, Florida, found the following weights (in pounds) of wildlife that had been killed by vehicles.

13.6 14.8 17.0 21.9 21.3 21.5 58.8 14.8 15.2 18.7 18.6 17.2
18.5 9.4 19.4 15.7 14.5 9.5 25.4 21.5 17.3 19.1 41.0 12.4
20.4 3.6 17.5 18.5 21.5 37.4

Part B. The following table shows the annual income, in dollars, and amount spent on vacation, in dollars, for a sample of 8 families

Income	41,200	55,000	28,400	35,600	65,900	98,400	72,000	57,600
Vacation	2,500	2,400	1,800	26,000	2,800	5,200	4,500	3,100

Instructional Items

Instructional Item 1

Jace surveyed his classmates on their favorite color and the results are shown in the table below. Determine if the data would be numerical (quantitative) or categorical (qualitative), univariate or bivariate, and create an appropriate display to represent the data.

Color	Frequency
Red	4
Orange	1
Yellow	2
Green	5
Blue	7
Purple	3

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

MA.912.DP.1.2

Benchmark

MA.912.DP.1.2	Interpret data distributions represented in various ways. State whether the data is numerical or categorical, whether it is univariate or bivariate and interpret the different components and quantities in the display.

Benchmark Clarifications:

Clarification 1: Within the Probability and Statistics course, instruction includes the use of spreadsheets and technology.

Connecting Benchmarks/Horizontal Alignment

- MA.912.DP.2
- MA.912.DP.3

Terms from the K-12 Glossary

- Bivariate Data
- Categorical Data
- Data

Vertical Alignment

Previous Benchmarks

- MA.6.DP.1
- MA.7.DP.1
- MA.8.DP.1

Next Benchmarks

Purpose and Instructional Strategies

In middle grades, students explored data and different types of graphical representations for univariate and bivariate data. Additionally, they worked with categorical and numerical data without formally stating what types of data they were using. In Algebra I, students interpreted the components of data displays for numerical and categorical data, both univariate and bivariate. In Mathematics for College Statistics, students continue with this exploration of being able to differentiate between univariate and bivariate data. Along with differentiating between numerical (quantitative) and categorical (qualitative) data.

- Instruction includes the use of terms such as quantitative data for numerical data and qualitative data for categorical data. Later statistics courses may use these terms interchangeably and students should have knowledge of these terms in order to avoid confusion.

- Instruction includes the use of spreadsheets and graphing calculators with statistical software.
- It is the intention of this benchmark to include cases where students must calculate measures of center (mean, median and mode) and measure of variation (standard deviation and interquartile range).
- The intention of this benchmark is to be taught in tandem with MA.912.DP.1.1. This is to reinforce students understanding of numerical/categorical (quantitative/qualitative) data, univariate/bivariate data sets and their displays.
    - The data displays for numerical (quantitative) univariate data includes histograms (grouped and ungrouped), relative frequency histograms (grouped and ungrouped), cumulative frequency histogram (grouped and ungrouped), stem-and-leaf plots, box plots and line plots (sometimes called dot plots).
    - The data displays for numerical (quantitative) bivariate data includes scatter plots and line graphs.
    - The data displays for categorical (qualitative) univariate data includes bar charts, circle graphs, frequency tables and relative frequency tables.
    - The data displays for categorical (qualitative) bivariate data includes side by side bar graphs, segmented bar graphs and two-way (contingency) tables
- This benchmark reinforces the importance of the use of questioning within instruction.
    - Is the data numerical (quantitative) or categorical (qualitative)?
    - Does this display univariate or bivariate data?
    - What do the different quantities within the data display mean in terms of the context of the situational data?
    - Which graphical display would be the most appropriate for the given data?
    - What is the shape, center and spread of the data? Are there any unusual data points?

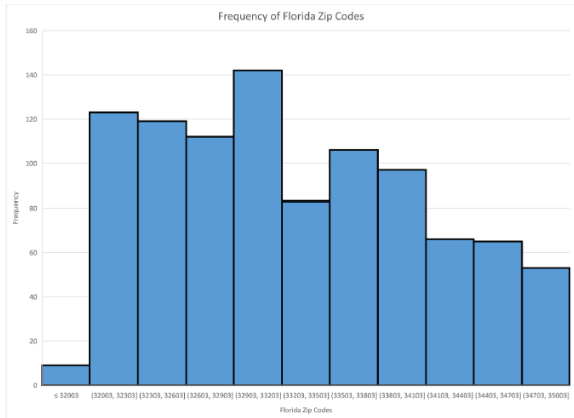
Common Misconceptions or Errors

- Students may not be able to properly distinguish between numerical (quantitative) and categorical (qualitative) data or between univariate and bivariate data.
- Students may misidentify and/or misinterpret the quantities in the various data displays.
- Students may not be able to distinguish between the measures of center (mean, median, mode) and measures of spread (standard deviation, interquartile range).
- Students may not completely grasp the effect of outliers on the data set; or incorrectly concludes a point is an outlier.
- Students may not be able to distinguish the differences between frequencies and relative frequencies or identify the condition that determines a conditional or relative frequency in a joint table.
- Students may believe that if the data consists of numbers then it must be numerical (quantitative) data.
    - For example, students may identify a list of zip codes as numerical (quantitative) data. This is not the case. Ask students whether there is a unit attached to these numbers or if it measure something. In this case, zip codes would be a categorical (qualitative) data set. So, it is not enough to just say numerical data is a data set that contains numbers.

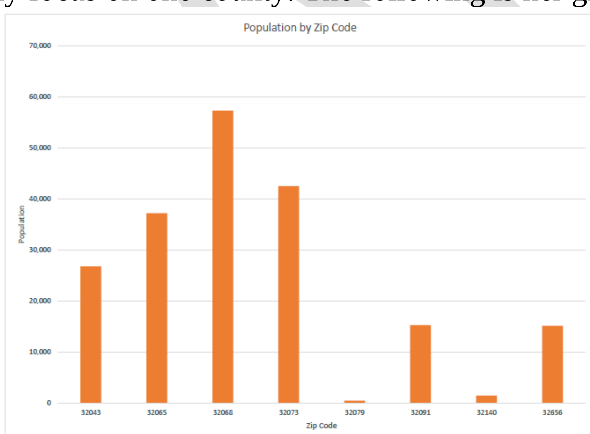
Instructional Tasks

Instructional Task 1 (MTR.2.1, MTR.4.1, MTR.7.1)

Jana went to Florida Demographics to look up data for an assignment. She found Florida zip codes and the corresponding population numbers for those zip codes (https://www.florida-demographics.com/zip_codes_by_population). She decided that the zip codes were numerical (quantitative) data and that she could create a histogram. The following is her histogram:



Sheila disagreed and said that zip codes are not numerical (quantitative) data and that making a histogram would not make sense. Sheila, instead, went to the same site and found each zip code and its population She said that zip codes were actually categorical (qualitative) data and that she could create a bar graph comparing their population. She decided not to use all of the zip codes and only focus on one county. The following is her graph:



Do you agree with either Jana or Sheila? Explain your reasoning.

If you agree with Jana, what information could be obtained from her graph? If you agree with Sheila, what information could be obtained from her graph? If you agreed with neither, what would have been a more appropriate graph and what information could be obtained from using such a visual display?

Instructional Items

Instructional Item 1

Shannon Marie Merrill from San Jose State University wrote a master's thesis titled "Individual Differences and Pet Ownership Status: Distinguishing Among Different Types of Pet Owners and Non-Owners." The following is her data as it relates to whether they were

male or female and whether they owned a dog, cat or both.
(https://scholarworks.sjsu.edu/cgi/viewcontent.cgi?article=7791&context=etd_theses)

	Dog	Cat	Dog and Cat
Male	69	13	80
Female	249	66	388

Part A. Is the data univariate or bivariate? Is it numerical (quantitative) or categorical (qualitative)?
Part B. How many people were surveyed that stated that they either owned dogs, cats or both?
Part C. Can we assume more females owned only dogs when compare to males? Why or why not?

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

MA.912.DP.1.3

Benchmark

MA.912.DP.1.3 Explain the difference between correlation and causation in the contexts of both numerical and categorical data.

Algebra I Example: There is a strong positive correlation between the number of Nobel prizes won by country and the per capita chocolate consumption by country. Does this mean that increased chocolate consumption in America will increase the United States of America's chances of a Nobel prize winner?

Connecting Benchmarks/Horizontal Alignment

- MA.912.DP.2.6, MA.912.DP.2.7
- MA.912.DP.5.7

Terms from the K-12 Glossary

- Association
- Bivariate data
- Categorical data

Vertical Alignment

Previous Benchmarks

- MA.8.DP.1.2, MA.8.DP.1.3

Next Benchmarks

Purpose and Instructional Strategies

In grade 8, students represented and investigated numerical bivariate data by constructing scatter plots and line graphs, and describing patterns of association. In Algebra I, students studied association between variables in bivariate data and learn that there is a difference between two variables being strongly associated and one of them having a causative effect on the other. In Mathematics for College Statistics, these ideas are extended by computing the correlation coefficient of a linear model. As well, students learn how to design statistical experiments that can show causation. In a college statistics course, students will continue to apply the ideas of linear regression and correlation.

- Instruction includes discussing confounding variables or lurking variables. A confounding variable may be defined as an extraneous variable that was not appropriately controlled. A lurking variable that is an unknown variable that was not accounted for in the study.

- Instruction includes making connections to surveys, experiments, observational studies, and the conclusions that can be drawn (MA.912.DP.5.6 and MA.912.DP.5.7).

Ideal experiment, but hard to achieve

Most observations

	Random Assignment	No Random Assignment	
Random Sampling	Causation inferred and generalized to the whole population	Correlation generalized to the whole population (no causation)	**Conclusions Generalized to Population**
No Random Sampling	Causation inferred, but only for the sample	Correlation only generalized for the sample	**Conclusions Not Generalized to Population**
	Causation	**Correlation**	

Most experiments

Bad observation due to limitations

- Instruction includes making connections by evaluating data-based arguments and reports from diverse media, which is covered in standard MA.912.DP.5.11.
- Correlation and causation are often misunderstood. It is important for students to understand their relationship. Causation and correlation can exist at the same time; however, correlation does not imply causation. Causation explicitly applies to cases where an action causes an outcome. Correlation is simply a relationship observed in bivariate data. One action may relate to the other, but that action does not necessarily cause the other to happen, because both of them may be the result of a third "hidden variable."
  - Causation is possible, but it is also possible that correlation occurs from a third variable.
    - For example, if one states, "On days when I drink coffee, I feel more productive," it may be that one feels more productive because of the caffeine (causation) or because they spent time in the coffee shop drinking coffee where there are fewer distractions (third variable). Since one cannot determine whether the causation or the third variable results in correlation, then causation is not confirmed.
  - Causation seems unlikely and a third variable seems likely
    - For example, there is a strong correlation between the number of Nobel prizes won by country and the per capita chocolate consumption by country. However, there are many possibilities a third variable, such as a strong economy, that can result in this correlation so causation can be ruled out.
  - Causation is likely because there is a reasonable explanation for the causation.
    - For example, if one states, "After I exercise, I feel physically exhausted," it is reasonable to consider this to be a cause-and-effect. Causation can be confirmed by the explanation that because one is purposefully pushing their body to physical exhaustion when doing exercise, the muscles used to exercise are exhausted (effect) after they exercise (cause).
  - When correlation is apparent in a bivariate data set, students are encouraged to seek a reasonable explanation that either identifies a hidden variable or a reasonable explanation for causation. Further investigation may be required to confirm or disconfirm causation.

Common Misconceptions or Errors

- Students may incorrectly believe correlations only applies to numerical data.
- Students may incorrectly assume that we can never infer a cause-and-effect relationship between variables.

Instructional Tasks

Instructional Task 1 (MTR.7.1)

Suppose we conduct an observational study by asking a random sample of 100 students two questions: 1) do you play an organized team sport? and 2) what is your current GPA? We find that students who play an organized team sport have a significantly higher GPA than those students who do not play an organized team sport.

Part A. What type of data will the first question yield?

Part B. What type of data will the second question yield?

Part C. Can we conclude that playing an organized team sport will cause a student's GPA to increase? Why or why not?

Part D. List any lurking variables that may exist.

Instructional Task 2 (MTR.7.1)

In a controlled experiment, 150 students planning to take the SAT in the near future are randomly assigned to one of three groups:

1) a control group in which students are not provided any additional SAT test preparation;

2) a treatment group in which students are provided with SAT test preparation in a large group classroom setting; and

3) a treatment group in which students are provided with SAT test preparation one-on-one with a personal tutor.

Suppose that we find that students who were provided with SAT test preparation one-on-one with a personal tutor scored significantly higher on the math portion of the test.

Part A. Would you infer there is an association or a cause-and-effect relationship between the type of test preparation and the student's math score on the SAT? Explain using complete sentences.

Part B. Suppose there was no significant difference in the reading SAT scores for students in the three groups. What is an appropriate conclusion?

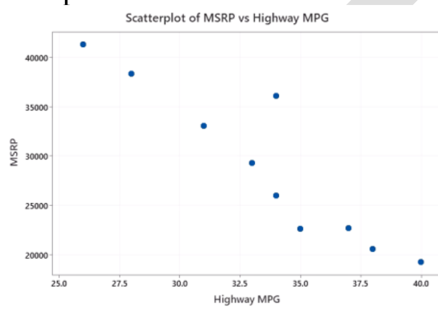
Instructional Items

Instructional Item 1

An analysis of grocery store sales and crime rates shows that as stores sell more hot chocolate the number of crimes in a city decrease. Can we can conclude that drinking hot chocolate leads to a decrease in the number of crimes?

Instructional Item 2

The scatter plot below shows 10 vehicles’ highway miles per gallon (MPG) and their suggested retail cost (MSRP – Manufacturer’s Suggested Retail Price); the trend is negative showing that as the highway miles per gallon increases the cost decreases. Write an appropriate conclusion using complete sentences.



**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

MA.912.DP.2 *Solve problems involving univariate and bivariate numerical data.*

MA.912.DP.2.1

Benchmark

MA.912.DP.2.1	For two or more sets of numerical univariate data, calculate and compare the appropriate measures of center and measures of variability, accounting for possible effects of outliers. Interpret any notable features of the shape of the data distribution.

Benchmark Clarifications:
Clarification 1: The measure of center is limited to mean and median. The measure of variation is limited to range, interquartile range, and standard deviation.
Clarification 2: Shape features include symmetry or skewness and clustering.
Clarification 3: Within the Probability and Statistics course, instruction includes the use of spreadsheets and technology.

Connecting Benchmarks/Horizontal Alignment

- MA.912.DP.1.1, MA.912.DP.1.2
- MA.912.DP.3.5

Terms from the K-12 Glossary

- Mean
- Median
- Measures of center
- Cluster (data)
- Data
- Interquartile range (IQR)
- Quartiles
- Standard deviation
- Range (of a data set)
- Measures of variability
- Outlier
- Histogram
- Box plot
- Line plot
- Stem-and-leaf plot

Vertical Alignment

Previous Benchmarks

- MA.6.DP.1
- MA.7.DP.1.1, MA.7.DP.1.2, MA.7.DP.1.5

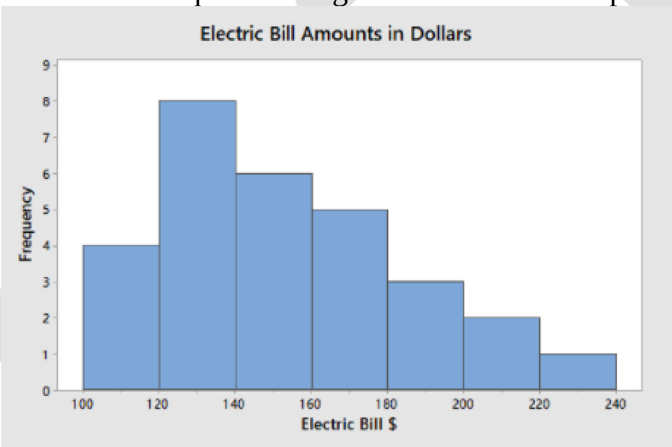
Next Benchmarks

Purpose and Instructional Strategies

In grades 6 and 7, students examined how to calculate and interpret the mean, median, mode and range of univariate numerical data. In grade 7, students compare numerical data sets and graphical representations. Additionally, students learned how to create histograms, box plots and line plots and interpret properties such as the center, spread and distribution of data, symmetry skewness, clustering, and outliers. In Mathematics for College Statistics, students continue exploring these ideas to use the shape of the distribution of data to determine the appropriate measures of center and variability and to make comparisons for two or more univariate numerical data sets.

- Instruction includes a review of mean, median, range and interquartile range (IQR) and a lesson on calculating standard deviation. Students should also learn to calculate outliers for univariate numerical data sets. Values that are two or more standard deviations from the mean are classified as outliers or "unusual" when data are symmetric. When data sets are skewed students should calculate the fences of $Q_1 - 1.5 * IQR$ and $Q_3 + 1.5 * IQR$ to determine outliers.

- Students should have practice using technology to make these calculations, but showing students how to use the formulas can reinforce the purpose of these values, especially standard deviation. It is recommended that students see how to calculate standard deviation with the formula using a smaller data set of five to seven numbers, as the calculation becomes tedious with larger data sets.
- Instruction includes creating histograms, boxplots, stem-and-leaf plots, and line plot using technology. Students should be reminded of identifying symmetry, skewness and outliers from these displays and how to determine the appropriate measures of center and variability.
  - For example, the histogram is right-skewed and the median is the better measure of center and the interquartile range is the measure of spread.



- Instruction includes how the mean and standard deviation are affected by skewness and outliers and why the median and IQR are better measures of center and variability when a data set is skewed and/or has outliers. The limitations of the range as a measure of variability should also be discussed.

Common Misconceptions or Errors

- Students may reverse left-skewed distributions and right-skewed distributions; emphasize skewness is in the direction of the longer tail.
- Students may believe that a distribution needs to be perfectly symmetric to utilize the mean and standard deviation; emphasize that the distribution should be approximately symmetric.
- Students may choose the value for the population standard deviation rather than the sample standard deviation when using technology.

Instructional Tasks

Instructional Task 1 (MTR.4.1, MTR.7.1)

In an observational study examining smoking habits and cholesterol levels, certain data were collected and published in the data and story library. The first data set shows the cholesterol levels of people who have smoked for at least 25 years. The second data set shows the cholesterol levels of people who smoked for no more than 5 years and quit. (https://dasl.datadescription.com/datafile/cholesterol-and-smoking/)

Smokers: 225, 221, 209, 284, 258, 216, 288, 250, 200, 209, 280, 225, 256, 243, 200, 213, 246, 225, 237, 232, 267, 232, 216, 216, 243, 200, 216, 271, 230, 309, 280, 217, 305, 287, 217, 246, 351, 200, 280, 209

Ex-Smokers: 250, 134, 300, 249, 213, 310, 175, 174, 328, 160, 188, 321, 213, 257, 292, 200, 271, 227, 238, 163, 263, 192, 242, 249, 242, 267, 243, 217, 267, 218, 217, 183, 228

Part A. Use technology to create an appropriate graphical representation for the smokers data and for the ex-smokers data.

Part B. Describe any notable features in the shapes of the two graphs. Determine if there are any outliers.

Part C. Calculate and interpret the appropriate measures of center and variability for the smokers data.

Part D. Calculate and interpret the appropriate measures of center and variability for the ex-smokers data.

Part E. Have students discuss their results by comparing the centers and variabilities for both groups and justify their reasoning.

Instructional Items

Instructional Item 1

The following test scores are collected from randomly sampled students in two different classes.

Class 1 test scores: 72, 75, 78, 78, 80, 82, 82, 85, 89, 90

Class 2 test scores: 42, 54, 72, 74, 77, 79, 82, 84, 88, 90

Calculate the mean, median, standard deviation, interquartile range, and range of the two data sets. Note which measure of center and which measure of variability is more appropriate for each class and explain why.

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

MA.912.DP.2.4

Benchmark

MA.912.DP.2.4	Fit a linear function to bivariate numerical data that suggests a linear association and interpret the slope and y-intercept of the model. Use the model to solve real-world problems in terms of the context of the data.

Benchmark Clarifications:
Clarification 1: Instruction includes fitting a linear function both informally and formally with the use of technology.
Clarification 2: Problems include making a prediction or extrapolation, inside and outside the range of the data, based on the equation of the line of fit.

Connecting Benchmarks/Horizontal Alignment

- MA.912.AR.2.5
- MA.912.F.1.2, MA.912.F.1.8
- MA.912.DP.1.1, MA.912.DP.1.2, MA.912.DP.1.3

Terms from the K-12 Glossary

- Association
- Bivariate data
- Line of fit
- Scatter plot
- Slope
- y-intercept

Vertical Alignment

Previous Benchmarks

- MA.8.AR.3.2, MA.8.AR.3.3, MA.8.AR.3.5
- MA.8.DP.1.1, MA.8.DP.1.2, MA.8.DP.1.3
- MA.912.AR.2.4

Next Benchmarks

Purpose and Instructional Strategies

In grade 8, students studied how to write a linear relationship and interpret the components from a written description, table, graph or equation. In addition, students constructed scatter plots, described patterns of association, and were able to informally fit a straight line to the data. In Algebra I, students continued this concept of graphing linear functions and finding lines of fit for data that could be represented with a linear function. In Mathematics for College Statistics, students continue to use the concept of the linear function in order to fit a linear equation to a set of bivariate numerical data that suggests a linear association from a scatter plot. Students also interpret the slope and the y-intercept of the linear function in the context of a real-world scenario; making predictions and extrapolations should also be addressed.

- Instruction addresses how to use technology to create scatter plots of numerical bivariate data. The “eyeball test” can be used to informally assess if there is an approximate linear association suggested by the data or if there is a clear nonlinear relationship.
- Instruction includes noting that (“y-hat”) denotes a prediction for a y-value and not the actual y-value.
- Students should have practice using technology to calculate the values for the y-intercept and the slope of the linear regression equation for numerical bivariate data that shows an approximate linear association. The regression equation is usually written in the form $\hat{y} = a + bx$ or $\widehat{f(x)} = a + bx$, where a is the y-intercept and b is the slope. This is different from the $y = mx + b$ form already familiar to students.

- Instruction includes the use of technology to graph the linear regression equation on the scatter plot to show how it is the line of best fit for the data.
- Instruction addresses making predictions within the observed domain of the independent variable as interpolation, and making predictions outside of the observed domain of the independent variable as extrapolation. Issues with knowing how the relationship continues outside the observed domain and why extrapolation can lead to illogical results should be communicated.
    - For example, suppose that the distance of domestic flight in miles (independent) and the cost of a domestic flight in dollars (dependent) show a linear association and a regression equation is calculated to be $\mathrm{f}(x) = 91.83 + 0.19x$. We could predict a flight of 156 miles would cost $91.83 + 0.19(156)$ which is equivalent to \$121.47. If 156 miles within the observed domain of the flight distances, this should be a fairly accurate prediction and interpolation.
- Instruction includes interpreting the slope and the y-intercept of the linear regression equation. The slope, b, can be interpreted as a prediction for the dependent variable as the independent variable increases by 1 unit (or for each additional unit). The y-intercept, a, can be interpreted as a prediction for the dependent variable when the independent variable is 0.
    - For example, suppose that the distance of domestic flight in miles (independent) and the cost of a domestic flight in dollars (dependent) shows a linear association and a regression equation is calculated to be $f(x) = 91.83 + 0.19x$. The slope could be interpreted as "the predicted cost of a flight will increase by \$0.19 for each additional mile." The y-intercept could be interpreted as "a flight that is 0 miles is predicted to cost \$91.83;" this illogical conclusion would be addressed as an extrapolation since a flight of 0 miles would cost \$0.

Common Misconceptions or Errors

- Students may reverse the values when interpreting the slope and the y-intercept.
- Students may confuse the independent and dependent variables.
- Students may believe that the regression equation should go through each data point on the scatter plot.
- Students may incorrectly think that the regression equation should go through the first and the last data point when the regression equation is plotted on the scatter plot.
- Students may have difficulty when determining if the association is actually best modeled with a linear function; methods are addressed in other benchmarks to help with clarification.

Instructional Tasks

Instructional Task 1 (MTR.2.1, MTR.7.1)

According to the Consumer Reports testing, 9 randomly selected motor vehicles had the following weights in pounds (lbs) and the corresponding highway miles per gallon (MPG).

*Data adapted from the Data and Story Library: https://dasl.datadescription.com/datafile/fuel-efficiency/

Weight (lbs)	2469	2606	2678	2822	3181	3306	3509	3648	3777
Highway MPG	34	33	33	29	25	26	26	24	19

Part A. Use technology to create a scatter plot of the data. Let weight be the independent variable and let MPG be the dependent variable. Does there appear to be a linear association?

Part B. Use technology to find the linear regression equation. Use complete sentences to interpret the slope and the y-intercept in the context of weight and miles per gallon.

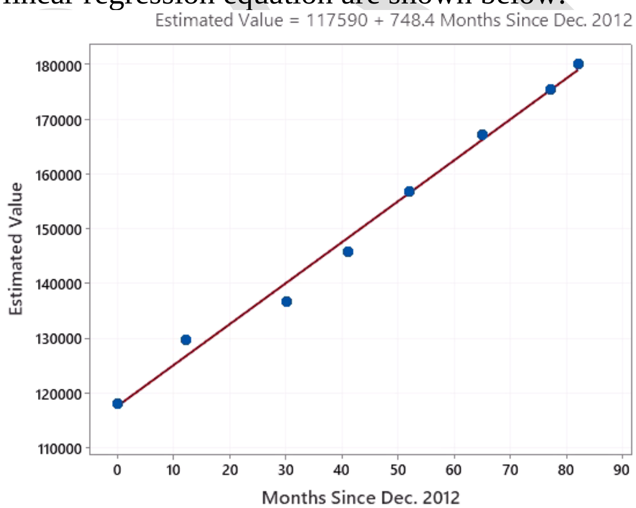
Part C. Use the regression equation to predict the highway miles per gallon for a vehicle that weighs 2,700 lbs. Do you believe that this would be an accurate prediction? Why or why not?

Part D. Use the regression equation to predict the highway miles per gallon for a vehicle that weighs 5,000 lbs. Do you believe that this would be an accurate prediction? Why or why not?

Instructional Items

Instructional Item 1

A homeowner tracks the estimate price of her home using a real estate app over several years. The data and linear regression equation are shown below.



Part A. Explain why there appears to be a linear association among the data points in the scatter plot.

Part B. Interpret the slope of the regression equation in the context of the problem. (Note the regression equation is $\hat{y} = 117590 + 748.4x$.)

Part C. Why should you be careful when predicting the estimated value of the home in December 2022, which is 120 months after December 2012?

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

MA.912.DP.2.5

Benchmark

MA.912.DP.2.5 Given a scatter plot that represents bivariate numerical data, assess the fit of a given linear function by plotting and analyzing residuals.

Benchmark Clarifications:
Clarification 1: Within the Algebra I course, instruction includes determining the number of positive and negative residuals; the largest and smallest residuals; and the connection between outliers in the data set and the corresponding residuals.

Connecting Benchmarks/Horizontal Alignment

- MA.912.DP.1.2

Terms from the K-12 Glossary

- Bivariate Data
- Line of Fit
- Scatter Plot

Vertical Alignment

Previous Benchmarks

- MA.8.DP.1.1, MA.8.DP.1.2, MA.8.DP.1.3

Next Benchmarks

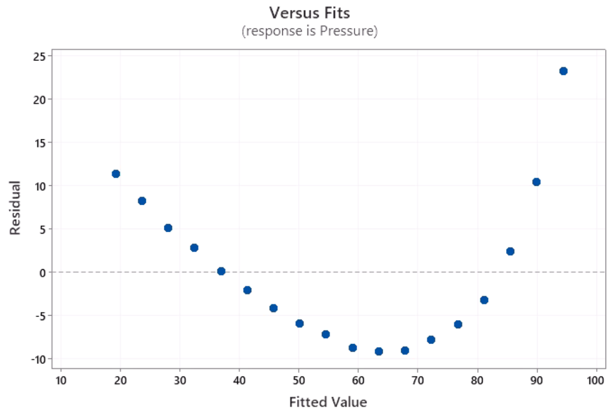
Purpose and Instructional Strategies

In grade 8, students constructed scatter plots, described patterns of association, and informally fit a straight line in scatter plots with linear associations. In Algebra I, students determined if residuals were positive or negative and if residuals were larger or smaller. In Mathematics for College Statistics, students expand upon these ideas by finding the linear regression equation for scatter plots of numerical bivariate data that suggest a linear association. With this benchmark students can assess the appropriateness of using a linear model in a more formal manner.

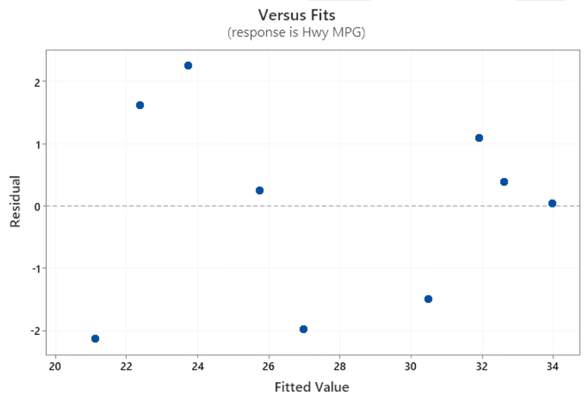
- Instruction includes using an actual data point and using the x-value to make a prediction for the y-value using the linear regression equation. The residual will be calculated from the actual y-value minus the predicted y-value. The residual should also be interpreted.
  - For example, suppose a data point for the number of minutes a student studied for a test and his actual grade on the test is represented by the data point $(35, 95)$ and the regression equation is $\widehat{T(x)} = 63.17 + .97x$ where $\widehat{T(x)}$ represents the predicted test score. A prediction for a student would be $63.17 + .97(35) = 97.12$. The residual would be $95 - 97.12 = -2.1$, indicating that the student scored around 2 points lower than predicted.
- Instruction includes the use of technology to plot residuals on the scatter plot with the plotted linear regression equation. Students should look at the number of data points above and below the line of best fit and should look at the size of the residuals to better understand the fit of the linear model to individual data points along with the full data set. Technology could also be used to calculate the corresponding standard deviation to assess which data points may be outliers.
- Instruction includes that a point with a large residual can be classified as an outlier.
- Technology may be used to create a residual plot with the original x-values and their corresponding residuals as y-values. Students should note that a clear pattern/association in the residual plot indicates that a linear model may not be the best fit, while a residual plot that has a weak association tends to indicate that a linear model is appropriate for the scenario.
  - Plot 1 below is a residual plot where the linear model does not seem to be appropriate, whereas plot 2 is a residual plot where the linear model does appear

to be a good fit.

Plot 1:



Plot 2:



Common Misconceptions or Errors

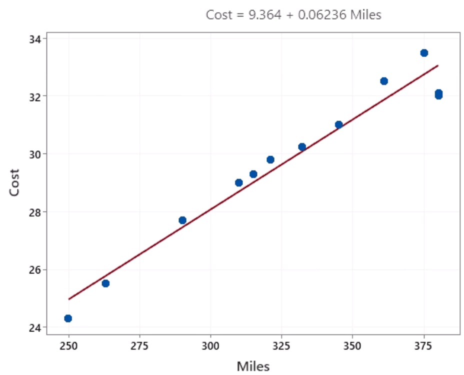
- When calculating residuals, students may reverse the order and find predicted y-value minus actual y-value instead of the correct form of actual y-value minus predicted y-value.
- Students may draw residuals horizontally instead of vertically.
- Students may try to create a pattern in the residual plot when there is no association.
- Students may incorrectly assume that a plotted regression equation and scatter plot with large residuals automatically means that a linear model is not a good fit for the data.
- Students may confuse a residual plot that is scattered with no obvious pattern to mean the linear model is not appropriate when in fact that is what is ideal.

Instructional Tasks

Instructional Task 1 (MTR.7.1)

Over the course of several months, a food delivery driver keeps track of how many miles he drives and how much he has to pay the next time he fills up his car. His data, scatter plot and corresponding linear regression equation are below.

Miles	361	321	380	290	345	332	250	263	380	375	310	315
Cost	32.50	29.80	32.00	27.70	31.00	30.25	24.30	25.50	32.10	33.50	29.00	29.30



$\widehat{C(x)} = 9.364 + 0.06236x$, where x is the number of miles driven.

Part A. Draw in the residuals on the scatter plot by constructing a vertical line from each point to the line of best fit. Which data point(s) appears to have the largest residual.

Part B. Calculate the residual value for the data points where the food delivery driver drove 250 miles and 321 miles.

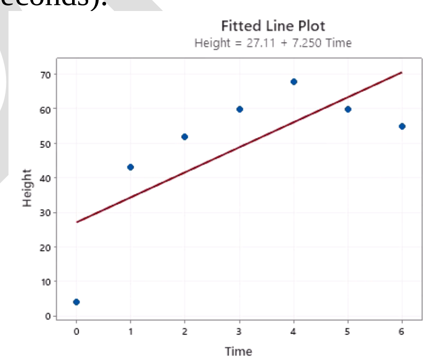
Part C. Use technology to create a residual plot, and using this plot, decide if the linear model is an appropriate choice for this data. Explain your reasoning.

Part D. Do any of the data points appear to be outliers? Explain your reasoning.

Instructional Items

Instructional Item 1

The scatter plot below is constructed from a set of bivariate data looking at the height of a drone in feet over time (in seconds).



Plot and analyze the residuals for this data. Determine if the linear function $\widehat{H(t)} = 27.11 + 7.25t$ is the best fit for this data. Explain your reasoning.

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

Benchmark

MA.912.DP.2.6	Given a scatter plot with a line of fit and residuals, determine the strength and direction of the correlation. Interpret strength and direction within a realworld context.

Benchmark Clarifications:

Clarification 1: Instruction focuses on determining the direction by analyzing the slope and informally determining the strength by analyzing the residuals.

Connecting Benchmarks/Horizontal Alignment

- MA.912.DP.1.2

Terms from the K-12 Glossary

- Association
- Bivariate data
- Line of fit
- Scatter plot

Vertical Alignment

Previous Benchmarks

- MA.8.DP.1.1, MA.8.DP.1.2, MA.8.DP.1.3

Next Benchmarks

Purpose and Instructional Strategies

In grade 8, students constructed scatter plots and informally fit a straight line into scatter plots with linear associations. Association in a scatter plot was described as positive/negative, linear/nonlinear and strong/weak. In Algebra I, students began determining strength, direction, line of fit and residuals. In Mathematics for College Statistics, students continue to describe the correlation, or association, among the data using the slope to assist in determining direction and the size of the residuals to aid in describing the strength.

- In MA.912.DP.2.7, students calculate the correlation coefficient to describe the strength and trend of the linear model in a more formal manner.
- Instruction addresses the sign of the slope and the trend of the scatter plot when determining the direction of the correlation.
- The size of residuals should be addressed in order to decide whether there is a strong, weak, or possibly moderate strength to the correlation.
- Instruction includes the use of real-world data.
- Instruction includes statistical applications or graphing utilities in order to help students visualize residuals.
- Instruction includes that a residual is calculated from the predicted value of the response variable subtracted from the actual value of the response variable and that residuals can be positive or negative.
- The residual is computed as, residual $= y - \hat{y}$. Some texts may use the variable e to note the residual or this can be abbreviated as "resid" (like most calculators).
- Instruction includes interpreting the strength and direction in the context of the scenario.
    - For example, if a scatter plot displaying data relating professional basketball teams' total salaries to total wins in a season shows a positive trend based on the slope and a strong correlation based on residuals, students could conclude that "there is a strong, positive relationship between total salaries and total wins; as the total salary for a team increase, the wins also tends to increase."

Common Misconceptions or Errors

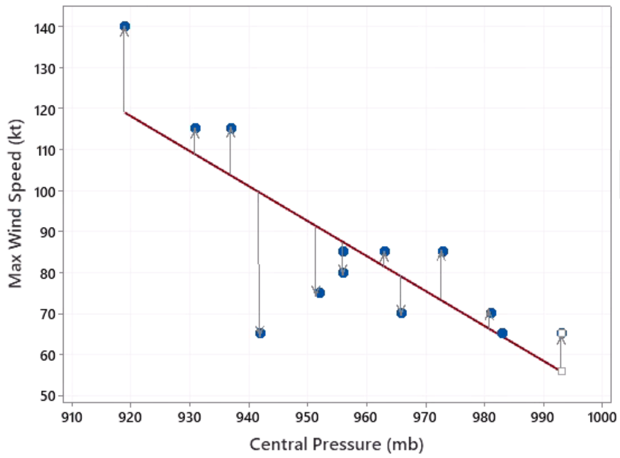
- Students may not be able to determine the strength in some scatter plots that are not clearly strong or clearly weak.

- Students may incorrectly assume that as the independent variable increase the dependent variable will also automatically increase. However, in scatter plots with a negative correlation, the dependent variable will decrease as the independent variable increases.
- Students may be confused by the meaning of a negative residual.

Instructional Tasks

Instructional Task 1 (MTR.4.1, MTR.7.1)

The strongest hurricanes from 2010 to 2019 were studied, and the central pressure (mb) and maximum wind speed (kt) were recorded and plotted in the scatter plot below. The linear regression equation is calculated to be $Predicted\ Max\ Wind\ Speed\ =\ 904.5\ -\ .85(Central\ Pressure)$.



Part A. Determine the direction of the scatter plot using the plotted line of fit.

Part B. How does the slope affirm the direction that was selected in part A?

Part C. Finish the sentence: As the central pressure during a hurricane increases, the maximum wind speed will typically __________.

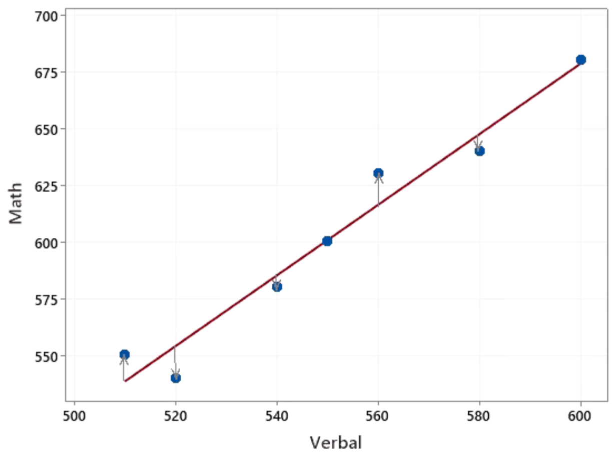
Part D. How would you describe the strength of the correlation in this scatter plot?

Part E. Which two values of central pressure have the largest residuals? With a partner, discuss how the removal of these points may or may not affect the strength of the correlation.

Instructional Items

Instructional Item 1

Data were collected on 7 students' SAT scores on the verbal and math sections. The results are in the scatter plot below; also plotted are the line of best fit and the residuals. The line of best fit equation is $\hat{y} = -255.4 + 1.556x$.



In a complete sentence, how would you describe the direction and strength of the correlation? How would you explain the relationship plotted above in the context of SAT scores on the verbal and math sections?

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

MA.912.DP.2.7

Benchmark

MA.912.DP.2.7 Compute the correlation coefficient of a linear model using technology. Interpret the strength and direction of the correlation coefficient.

Connecting Benchmarks/Horizontal Alignment

- MA.912.DP.1.2

Terms from the K-12 Glossary

- Association
- Bivariate data
- Line of fit
- Scatter plot

Vertical Alignment

Previous Benchmarks

- MA.8.DP.1.1, MA.8.DP.1.2, MA.8.DP.1.3

Next Benchmarks

Purpose and Instructional Strategies

In grade 8, students constructed scatter plots and informally fit a straight line into scatter plots with linear associations. Association in a scatter plot was described as positive/negative, linear/nonlinear and strong/weak. In Mathematics for College Statistics, students continue to describe the correlation, or association, among the numerical bivariate data using the correlation coefficient to assist in determining direction and the size of the residuals to aid in describing the strength.

- In MA.912.DP.2.6, students use the scatter plot, residuals and slope to describe the strength and trend of the linear model. This benchmark expands on those ideas using the correlation coefficient.
- Instruction relies on using technology to calculate the correlation coefficient. Using a formula can be quite tedious, and may rely upon calculations that students are not familiar with at this point in their learning.
- This benchmark is only meant to address linear models.
- Instruction emphasizes the correlation coefficient, or r, is always between -1 and 1. A perfect positive linear association has a correlation coefficient of 1; a perfect negative linear association has correlation coefficient of -1. Weaker relationships have correlations that are closer to 0.
- Instruction emphasizes that the sign of the correlation coefficient corresponds to the direction/trend of the scatter plot and the sign of the slope. It should also be noted that the sign of the correlation coefficient does not affect the strength of the correlation.
- While there are not an agreed upon set of boundaries when classifying the strength of a correlation, it may be helpful to set up guidelines for students.
    - For example, a statistics course may generally teach that if $|r| \geq .80$ then the relationship is strong, if $.50 \leq |r| \leq .79$ then the relationship is moderately strong, and if $|r| \leq .49$ then the relationship is weak.
- Instruction includes describing the direction of the correlation as positive or negative and the strength of the correlation as strong, possibly moderate or weak.
- It should be mentioned that the formal name of this computation is Pearson's correlation coefficient, and it is used to measure the strength and direction of linear models. However, there are other calculations that exist for measuring the strength of nonlinear models.

Common Misconceptions or Errors

- Students may incorrectly assume that the sign of the correlation has an effect on the strength.
- Students may have issues comparing the decimal values of correlation coefficients.
- Students may have trouble rounding decimals or rounding to the correct decimal place.

Instructional Tasks

Instructional Task 1 (MTR.6.1, MTR.7.1)

The strongest hurricanes from 2010 to 2019 were studied, and the central pressure (mb) and maximum wind speed (kt) were recorded and are given in the table below.

Central Pressure (mb)	952	966	942	973	981	963	937	931	983	956	919	993	956
Max Wind Speed (kt)	75	70	65	85	70	85	115	115	65	80	140	65	85

Part A. Use technology to create a scatter plot, and use the graph to estimate the correlation coefficient.
Part B. Use technology to calculate the correlation coefficient.
Part C. Describe the direction of the correlation coefficient.
Part D. Describe the strength of the correlation coefficient.
Part E. How does your estimate in Part A compare to the calculation in Part B?
Part F. Use technology to find the line of best fit. Do you feel the linear model is appropriate based on the graph and the correlation coefficient? Justify your reasoning.

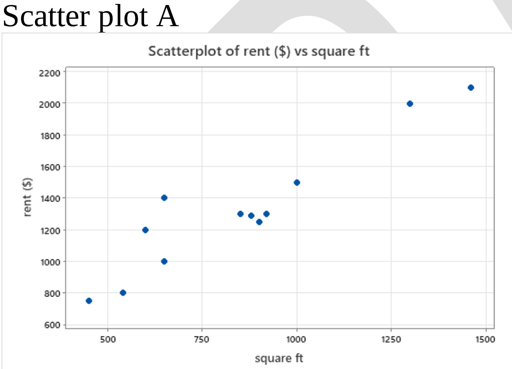
Instructional Items

Instructional Item 1

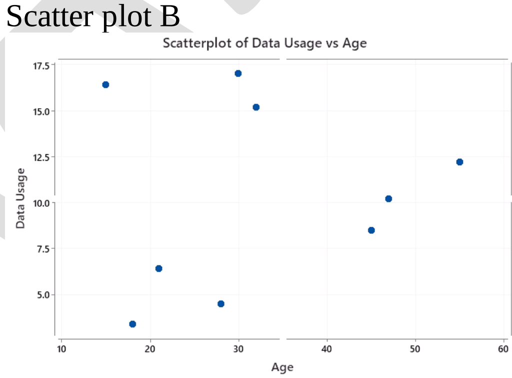
Match the correlation coefficient with the appropriate scatter plot.

- $r = -1$
- $r = .94$
- $r = -.75$
- $r = .11$

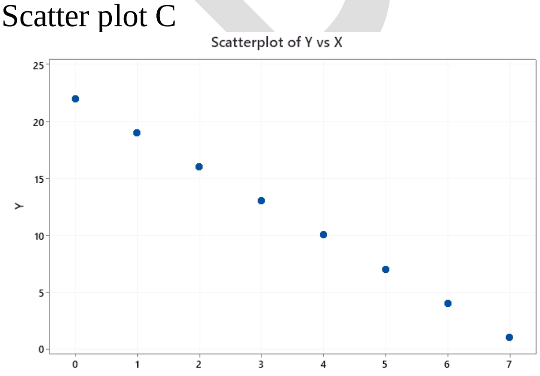
Scatter plot A



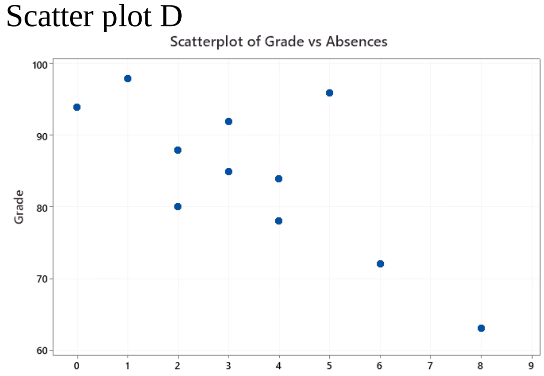
Scatter plot B



Scatter plot C



Scatter plot D



Instructional Item 2

Use technology to calculate the correlation coefficient of the data set. Interpret the direction and strength.

Square Footage vs Rent in Chicago	
Square Feet	**Rent ($)**
650	1000
540	800
600	1200
1300	2000
900	1250
850	1300
650	1400
450	750
1460	2100
1000	1500
920	1300
880	1290

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

MA.912.DP.2.9

Benchmark

MA.912.DP.2.9 Fit an exponential function to bivariate numerical data that suggests an exponential association. Use the model to solve real-world problems in terms of the context of the data.

Benchmark Clarifications:

Clarification 1: Instruction focuses on determining whether an exponential model is appropriate by taking the logarithm of the dependent variable using spreadsheets and other technology.

Clarification 2: Instruction includes determining whether the transformed scatter plot has an appropriate line of best fit, and interpreting the y-intercept and slope of the line of best fit.

Clarification 3: Problems include making a prediction or extrapolation, inside and outside the range of the data, based on the equation of the line of fit.

Connecting Benchmarks/Horizontal Alignment

- MA.912.AR.2.5
- MA.912.AR.5.7
- MA.912.F.1.2, MA.912.F.1.8
- MA.912.DP.1.1

Terms from the K-12 Glossary

- Bivariate data
- Exponential function
- Intercept
- Line of fit
- Scatter plot
- Slope
- y-intercept

Vertical Alignment

Previous Benchmarks	Next Benchmarks
• MA.912.AR.5.3, MA.912.AR.5.4, MA.912.AR.5.6 • MA.912.F.1.1, MA.912.F.1.6	

Purpose and Instructional Strategies

In Algebra I, students learned to identify, describe and graph exponential functions. As well, before students approach this benchmark in Mathematics for College Statistics, they should have covered MA.912.DP.1.4 where they fit a linear model to bivariate numerical data that display a linear trend. The purpose of this benchmark is to address how to approach bivariate numerical data that display a nonlinear trend, specifically an exponential trend. Depending on a student's previous coursework, they may be working with logarithmic functions for the first time.

- Instruction includes a review of the properties of exponential and a review or introduction to logarithmic functions.
- Instruction includes the usage of technology for graphing and calculations.
- Instruction includes real-world data where possible.
- Instruction demonstrates that scatter plots that display and exponential pattern require the response variable data to be transformed by taking the natural logarithm of each dependent value. If an exponential model is appropriate, the transformed data of the original explanatory/independent values and the transformed response/dependent values should display a linear trend when plotted in a scatter plot.
- Technology is also used to graph the linear regression equation on the scatter plot to show how it is the line of best fit for the data.
- Instruction includes using both logarithms (log) and natural logarithms (ln) to transform data.
- Technology should be used to calculate the values for the y-intercept and the slope of the linear regression equation for transformed numeric bivariate data that shows an approximate linear association. The regression equation is usually written in the form $\hat{y} = a + bx$ or $f(x) = a + bx$, however, for data that originally display an exponential pattern, the form is $ln(\hat{y}) = a + bx$ since the y-variable was transformed.
- Instruction includes interpreting the slope of the linear regression equation from the transformed data. The concept of the slope applies to the transformed data, however, the original data does not have a slope as it would be nonlinear. The slope would be interpreted as a prediction for the natural log of the dependent variable (the $ln(y)$) as the independent variable increases by 1.
- Instruction includes interpreting the y-intercept of the linear regression equation from the transformed data. The y-intercept, a, can be interpreted as a prediction for the natural log of the dependent variable when the independent variable is 0.
- Instruction addresses making predictions within the observed domain of the independent variable as interpolation, and making predictions outside of the observed domain of the independent variable as extrapolation. Issues with knowing how the relationship continues outside the observed domain and why extrapolation can lead to illogical results should be communicated.

Common Misconceptions or Errors

- Students may confuse the independent and dependent variables.
- Students may be unfamiliar with transforming the data using the natural log.
- Students may transform the independent variable rather than the dependent variable.
- Students may have difficulty interpreting the slope and the y-intercept of the regression equation generated from the transformed data.
- Students may have difficulty converting back to the original context after the data are transformed.
- Students may mix up log, ln and their inverse operations.
- Students may believe that the regression equation should go through each data point on the scatter plot.
- Students may incorrectly think that the regression equation should go through the first and the last data point when the regression equation is plotted on the scatter plot.
- Students may have difficulty with determining if the association is actually best modeled with a linear function; methods are addressed in other standards to help with clarification.

Instructional Tasks

Instructional Task 1 (MTR.6.1, MTR.7.1)

In the 17th century, scientist Robert Boyle devised and experiment in which he discovered that the volume of a gas decreases and pressure increases. He used a container shaped like a U and poured mercury into the tube. He measured the height of the mercury on one side of the tube and the resulting pressure on the other side. Some of the data the Boyle derived is in the table below.

(Source: https://journals.physiology.org/doi/full/10.1152/jappl.1999.87.4.1543)

Height (in)	24	26	28	30	32	34	36	38	40	42	44	46	48
Pressure (Kpa)	58.8	54.3	50.3	47.5	44.2	41.6	39.3	37	35.3	33.2	31.9	30.6	29.1

Part A. Use technology to create a scatter plot of the data. Let height be the independent variable and let pressure be the dependent variable. Does there appear to be a linear association? How would you describe the shape of the plot?

Part B. Calculate the natural log of each value for the pressure. Use technology to create a scatter plot of the data. Let height be the independent variable again and let $ln(pressure)$ be the dependent variable. The shape of the plot should be linear. Is an exponential model appropriate for the original data set?

Part C. Use technology to find the linear regression equation of the data used in Part B. Use complete sentences to interpret the slope and the y-intercept in the context of height and pressure.

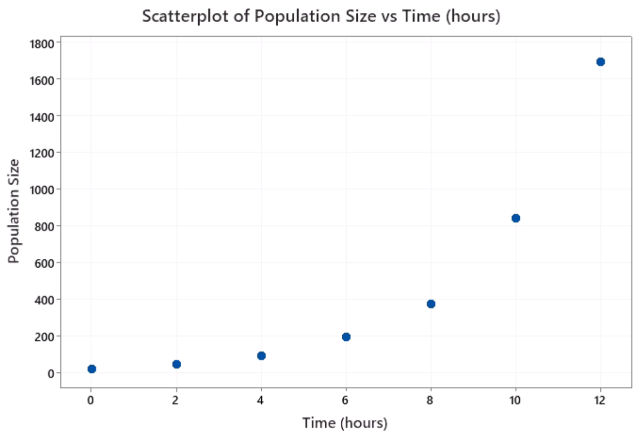
Part D. Use the linear regression equation to predict the pressure if the mercury is poured to a height of 35 inches. Do you believe that this would be an accurate prediction? Why or why not?

Part E. Use the regression equation to predict pressure if the mercury is poured to a height of 55 inches. Do you believe that this would be an accurate prediction? Why or why not?

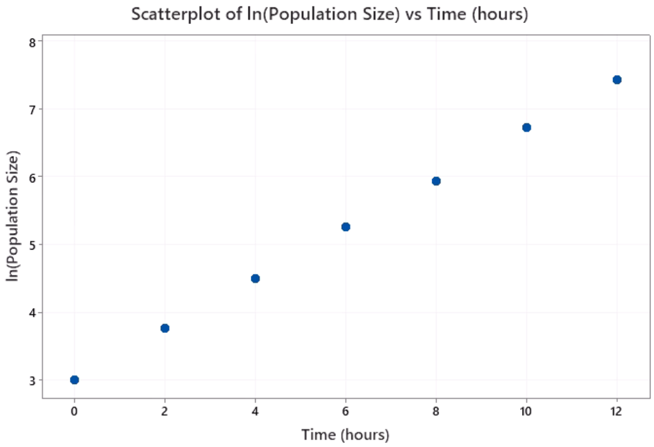
Instructional Items

Instructional Item 1

A scientist is measuring the population growth of fast-growing bacteria. He measures the number of bacteria in a petri dish every two hours over the course of a day. The initial scatter plot is below.



Part A. Noting that the shape is nonlinear, the scientist takes the logarithm of the dependent variable and graphs the transformed scatter plot, which is below.



How can we determine if the exponential model is appropriate?

Part B. The linear regression equation is $ln(population\ size) = 3.014 + 0.369(time)$. Interpret the slope and y-intercept in the context of the size of the bacteria population over time.

Part C. Should the scientist attempt to use the linear regression model to predict the population size 36 hours after the experiment has started? Elaborate.

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

MA.912.DP.3 *Solve problems involving categorical data.*

MA.912.DP.3.1

Benchmark

MA.912.DP.3.1 Construct a two-way frequency table summarizing bivariate categorical data. Interpret joint and marginal frequencies and determine possible associations in terms of a real-world context.

Algebra I Example: Complete the frequency table below.

	Has an A in math	Doesn't have an A in math	Total
Plays an instrument	20		90
Doesn't play an instrument	20		
Total			350

Using the information in the table, it is possible to determine that the second column contains the numbers 70 and 240. This means that there are 70 students who play an instrument but do not have an A in math and the total number of students who play an instrument is 90. The ratio of the joint frequencies in the first column is 1 to 1 and the ratio in the second column is 7 to 24, indicating a strong positive association between playing an instrument and getting an A in math.

Connecting Benchmarks/Horizontal Alignment

- MA.912.DP.1.1, MA.912.DP.1.2
- MA.912.DP.4.1, MA.912.DP.4.2, MA.912.DP.4.3, MA.912.DP.4.4, MA.912.DP.4.5, MA.912.DP.4.6

Terms from the K-12 Glossary

- Bivariate data
- Categorical data
- Joint frequency
- Joint relative frequency

Vertical Alignment

Previous Benchmarks

- MA.7.DP.1
- MA.8.DP.1
- MA.912.DP.2

Next Benchmarks

Purpose and Instructional Strategies

In grade 8, students began their exploration of bivariate data. Students focused on informally finding the line of fit. In Algebra I, students studied bivariate categorical data and displayed it in tables showing joint frequencies and marginal frequencies. In Mathematics for College Statistics, students continue this exploration with the use of real-world data.

- Instruction includes the connection to MA.912.DP.1.1 where students work with categorical bivariate data and display it in tables. A two-way frequency table is just a way to display frequencies jointly for two categories.
- In later courses, students may hear the term for two-way tables as contingency tables. Instruction includes the use of both terms interchangeably so that students are familiar with them.
- Students are expected to interpret the joint and marginal frequencies and therefore must

know the differences between these two.

- **Marginal frequencies**
  Entries in the "Total" row and "Total" columns are called marginal frequencies. The areas highlighted in yellow in the table below represents the marginal frequencies.

	Ride the Bus	Do Not Ride the Bus	Total
Boys	7	5	12
Girls	2	7	9
Total	9	12	21

- **Joint frequencies**
  Entries in the body of the table represent joint frequencies. The area highlighted in yellow in the table represents the joint frequencies.

	Ride the Bus	Do Not Ride the Bus	Total
Boys	7	5	12
Girls	2	7	9
Total	9	12	21

- Instruction includes students creating two-way (contingency) tables from data that has either been collected or researched. Students can then analyze the data by either examining joint relative frequencies or marginal relative frequencies.
- Instruction includes determining possible association. Two categorical variables are associated if the row of conditional relative frequencies (or column of relative frequencies) are different for the rows (or columns) of the table. Evidence of an association is strongest when the conditional relative frequencies are quite different. If the conditional relative frequencies are nearly equal for all categories, then there is probably not an association between variables.
- Instruction includes determining if variables are independent. If the conditional relative frequencies are nearly equal for all categories, there may be no association between the variables and they are said to be independent.
- Instruction includes experience with the Simpson's Paradox. Simpson's Paradox is an association or comparison that holds for all of several groups can reverse direction when the data are combined to form a single group.
  - For example, a university was concerned with the number of females who were being accepted into their graduate programs. Of the 384 men who applied to the school, 285 were admitted which is an admission rate of 74.2%. Of the 435 women who applied to the school, 126 were admitted which is an admission rate of 29.0%. This may suggest there is a bias where men are favored in graduate admissions. However, upon closer inspection of the graduate programs they offered, the following was discovered:

	Men			**Women**		
	Accepted	**Applied**	**Rate**	**Accepted**	**Applied**	**Rate**
Education	9	40	22.5%	45	127	35.4%
Physics	116	121	95.9%	5	5	100.0%
Nursing	7	60	11.7%	59	285	20.7%
Mathematics	153	163	93.9%	17	18	94.4%
Total	285	384	74.2%	126	435	29.0%

When looking at the individual programs and comparing the acceptance rate of women versus that of men, they noticed that more women were being accepted into these programs then men were. This is an example of a Simpson's Paradox. When the data is combined into a single group, we get one result. When the data is divided into several groups, the results are reversed.

Common Misconceptions or Errors

- Students may have difficulty when completing the table based on the given data.
- Students may not distinguish the differences between marginal and joint frequencies.
- Students may make errors when identifying the relationship and possible associations in the data in terms of the given context.
- Students may find the relative joint frequencies of the column only and not realize that to find the relative joint frequency depends on what is being analyzed and could be based on the row.

Instructional Tasks

Instructional Task 1 (MTR.3.1, MTR.7.1)

The U.S. Department of Transportation Federal Highway Administration performed a National Household Travel Survey. Within that survey, they looked at trends in the annual number (millions) of person trip by mode of transportation. The following is the data from 2017 (https://nhts.ornl.gov/assets/2017_nhts_summary_travel_trends.pdf):

- *Private Vehicle*: to and from work: 56,981; work-related business: 4,844; shopping and errands: 126,268; school or church: 28,427; social and recreational: 78,890; other: 10,988
- *Public Transit*: to and from work: 3,537; work-related business: 208; shopping and errands: 2,586; school or church: 1,009; social and recreational: 1,618; other: 487
- *Walk*: to and from work: 2,523; work-related business: 510; shopping and errands: 11,496; school or church: 4,146; social and recreational: 18,483; other: 1,790
- *Other*: to and from work: 1,540; work-related business: 486; shopping and errands: 2,404; school and church: 6,721; social and recreational: 3,330; other: 1,873

Part A. Create a contingency table using the data provided.

Part B. What is the joint relative frequency of people who used public transit traveled for work related business?

Part C. What is the marginal relative frequency of those who traveled for shopping and errands?

Part D. What is the joint relative frequency of people who traveled to and from work by using a private vehicle?

Part E. Which mode of transportation was most popular as it relates to social and recreational trips?

Part F. Does the data show an association between using a personal vehicle and going to school or church? Justify your reasoning.

Instructional Items

Instructional Item 1

The results from a survey about whether students (male or female) at a university were from England, Wales or Scotland is summarized in the following two-way table:

	Male	**Female**	**Total**
Scotland	34	16	
Wales	40	8	
England	2	0	
Total			

Part A. Complete the table with the totals.
Part B. What proportion of the students are women?
Part C. What proportion of students are from Wales?
Part D. Is it fair to say that more students from Wales are women? Explain your reasoning.

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

MA.912.DP.3.2

Benchmark

MA.912.DP.3.2 Given marginal and conditional relative frequencies, construct a two-way relative frequency table summarizing categorical bivariate data.

Algebra I Example: A study shows that 9% of the population have diabetes and 91% do not. The study also shows that 95% of the people who do not have diabetes, test negative on a diabetes test while 80% who do have diabetes, test positive. Based on the given information, the following relative frequency table can be constructed.

	Positive	Negative	Total
Has diabetes	7.2%	1.8%	9%
Doesn't have diabetes	4.55%	86.45%	91%

Benchmark Clarifications:
Clarification 1: Construction includes cases where not all frequencies are given but enough are provided to be able to construct a two-way relative frequency table.
Clarification 2: Instruction includes the use of a tree diagram when calculating relative frequencies to construct tables.

Connecting Benchmarks/Horizontal Alignment

- MA.912.FL.1.1
- MA.912.DP.1.1
- MA.912.DP.4.1, MA.912.DP.4.2, MA.912.DP.4.3, MA.912.DP.4.4

Terms from the K-12 Glossary

- Categorical data
- Bivariate data
- Conditional relative frequency
- Joint relative frequency

Vertical Alignment

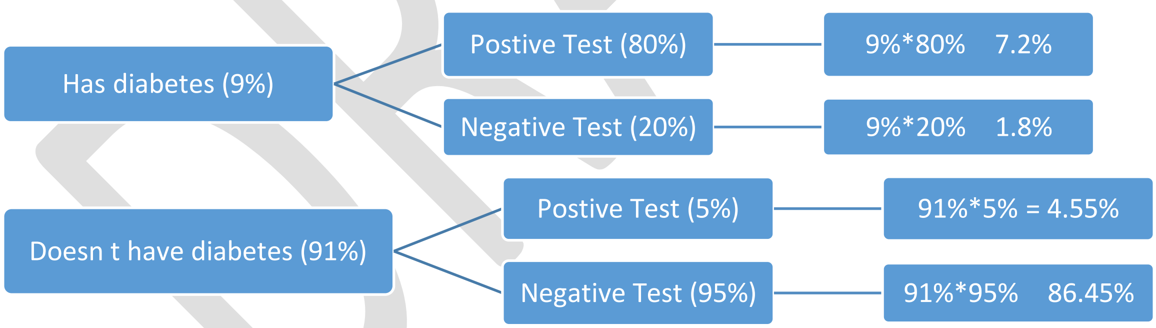
Previous Benchmarks
- MA.7.DP.1.5
- MA.8.DP.2.1

Next Benchmarks

Purpose and Instructional Strategies

In grades 7 and 8, students have not worked specifically with two-way tables, however, they have examined sample spaces and utilized tree diagrams to create a sample space. In Algebra I, students began to explore two-way tables. In Mathematics for College Statistics, students work with bivariate numerical data and bivariate categorical data. When analyzing bivariate categorical data, students should utilize two-way tables to summarize results and to determine possible associations.

- Instruction of constructing relative frequency two-way tables is interwoven with instruction on two-way tables using frequencies. A two-way table can also be referred to as a contingency table or joint frequency table.
- Instruction relates the relative frequencies to the probabilities for the events summarized in the table and that the sum of the joint relative frequencies should add to be 100%.
- As noted in the clarifications, students should be given some of the values and use the properties of two-way tables to find the missing values. Emphasize that the joint relative frequencies can be added vertically and horizontally to find the total row and total column.
- As noted in the clarifications, tree diagrams can be used as an alternative display. One of the categorical variables is used for the initial separation and then the other categorical variable is used for a secondary separation.
    - For example, the two-way table from the example in the benchmark would relate to the tree diagram below.



Common Misconceptions or Errors

- Students may try to input a given conditional relative frequency in place of the joint relative frequency in the two-way table. Constructing the tree diagram can help with this.
- Students may leave off the total row or total column in a contingency table.
- Students may have difficulty when relative frequencies are given in decimal form.
- Students may be unable to correctly construct a two-way table if given a tree-diagram especially because the middle branch of percentages does not appear in the table at all.

Instructional Tasks

Instructional Task 1 (MTR.2.1, MTR.4.1, MTR.7.1)

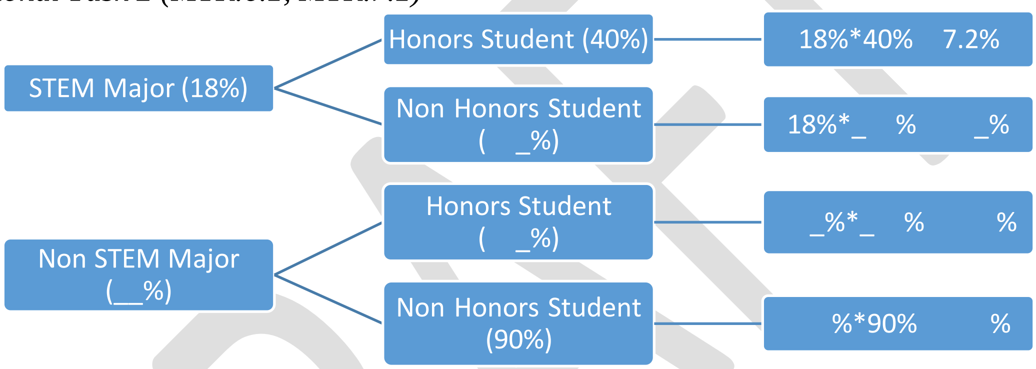
Suppose that 3% of Olympians have taken a banned performance-enhancing substance. If an Olympian has used one of these substances, he or she has a 95% chance of receiving a positive drug test. If an Olympian has not taken one of these substances, he or she has a 2% chance of receiving a positive drug test.

Part A. Construct either the corresponding relative frequency two-way table or the tree diagram for the scenario.

Part B. Using your two-way table, what is the relative frequency of a false positive – a positive drug test for someone who has not used a banned performance-enhancing substance?

Part C. Have a class discussion about the benefits or drawbacks of using the different diagrams from Part A.

Instructional Task 2 (MTR.6.1, MTR.7.1)



The dean of science and math at a Florida college looks at the relative frequencies of students at her campus who are science, technology, engineering and math (STEM) majors; who are non-STEM students; who are honors students; and who are non-honors students. Some of the relative frequencies from her findings are in the tree diagram above.

Part A. Fill in the missing components of the tree diagram. Do your results make sense? Justify your reasoning.

Part B. Use the tree diagram to construct a relative frequency table for this information.

Part C. What are the two categorical variables being studied?

Instructional Items

Instructional Item 1

A study shows that about 4% of men develop the skin cancer melanoma in their lifetime. Current medical testing is not always completely accurate. In cases where a man actually had melanoma, his test result came back positive about 92% of the time. In a case where a man did not actually have melanoma, his test came back positive about 1% of the time. Complete the table.

	Positive	Negative	Total
Has melanoma	3.68%		4%
Doesn’t have melanoma	0.96%		
Total			100%

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

MA.912.DP.3.5

Benchmark

MA.912.DP.3.5 Solve real-world problems involving univariate and bivariate categorical data.

Benchmark Clarifications:
Clarification 1: Instruction focuses on the connection to probability.
Clarification 2: Instruction includes calculating joint relative frequencies or conditional relative frequencies using tree diagrams.
Clarification 3: Graphical representations include frequency tables, relative frequency tables, circle graphs and segmented bar graphs.

Connecting Benchmarks/Horizontal Alignment

- MA.912.FL.1.1
- MA.912.DP.1.1, MA.912.DP.1.2
- MA.912.DP.4.2, MA.912.DP.4.3, MA.912.DP.4.4, MA.912.4.5

Terms from the K-12 Glossary

- Categorical data
- Bivariate data
- Bar graph
- Circle graphs
- Joint relative frequency
- Conditional relative frequency
- Frequency table

Vertical Alignment

Previous Benchmarks

- MA.7.DP.1.4, MA.7.DP.1.5
- MA.7.DP.2.2

Next Benchmarks

Purpose and Instructional Strategies

In grade 7, students worked with categorical data to create graphical displays. In addition, students also began to develop an understanding of probabilities in grade 7, and continued this further in grade 8. In Mathematics for College Statistics, students associate the relative frequencies seen in tables and graphs with probabilities to solve real-world problems.

- Instruction includes reinforcing that univariate categorical data can be displayed with circle graphs and bar graphs and that the relative frequencies associated with these graphs can be viewed as probabilities.
- Instruction includes reinforcing that bivariate categorical data can be summarized with segmented bar graphs, tree diagrams, frequency two-way tables, and relative frequency two-way tables. Various probabilities can be calculated from these displays.
- Instruction makes the connection to joint relative frequencies and conditional relative frequencies.
- When connecting instruction to probability, emphasize that probabilities can be represented as fractions, decimals or percentages. Students should realize that calculations from categorical data are typically probabilities/percentages/fractions since this data type is qualitative.
- Instruction includes the use of technology to create graphical displays of data.
- Instruction includes a discussion on false positives and false negatives. A false positive occurs when a test result comes back as positive when the result should have been negative. A false negative occurs when a test result comes back as negative when the

result should have been positive.

Common Misconceptions or Errors

- When converting a fraction to a decimal, students may divide the denominator by the numerator.
- Students may not fully fill in a frequency two-table or a relative frequency two-way table, which would lead to incorrect calculations.

Instructional Tasks

Instructional Task 1 (MTR.7.1)

A study shows that about 2.5% of women develop the skin cancer melanoma in their lifetime. Current medical testing is not always completely accurate. In cases where a woman actually had melanoma, her test result came back positive about 94% of the time. In a case where a woman did not actually have melanoma, her test came back positive about 2% of the time. A dermatologist’s office is interested to know whether their rates are similar to the rates of the study. The results are in the two-way table below.

	Positive	Negative	Total
Has melanoma	56	4	60
Doesn’t have melanoma	6	244	250
Total	62	248	310

Part A. Is the data being summarized univariate or bivariate?

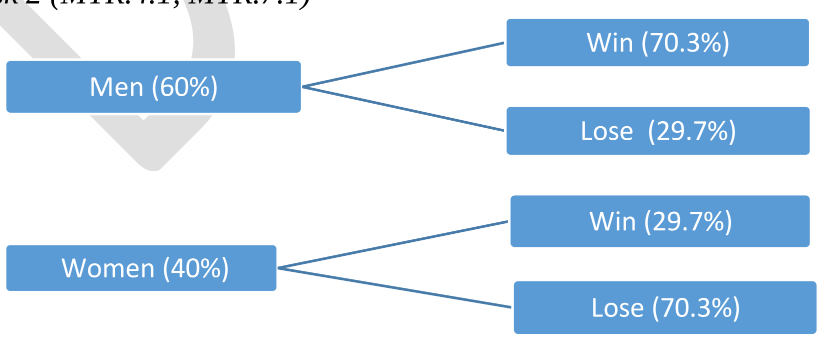
Part B. What is the conditional probability that a woman had a positive test given that she actually had melanoma? Is this similar to the result of the aforementioned study?

Part C. What is the joint probability that a woman had a positive test and did not have actually have melanoma? Is this a false positive or a false negative?

Part D. What is the conditional probability that a woman had a positive test given that she did not actually have melanoma? Is this similar to the result of the aforementioned study?

Part E. What is the joint probability that a woman had a negative test and actually had melanoma? Is this a false positive or a false negative?

Instructional Task 2 (MTR.4.1, MTR.7.1)



A 2014 article by Slate looked at the sex of contestants on the gameshow Jeopardy! and whether or not they won. The statistics provided by that article are in the tree diagram above.

Part A. What is the probability that a woman appeared on an episode of Jeopardy! and won?

Part B. What is the probability that a man appeared on an episode of Jeopardy! and won?

Part C. Given that the contestant was a man, what is the probability that he won on an episode of Jeopardy!?

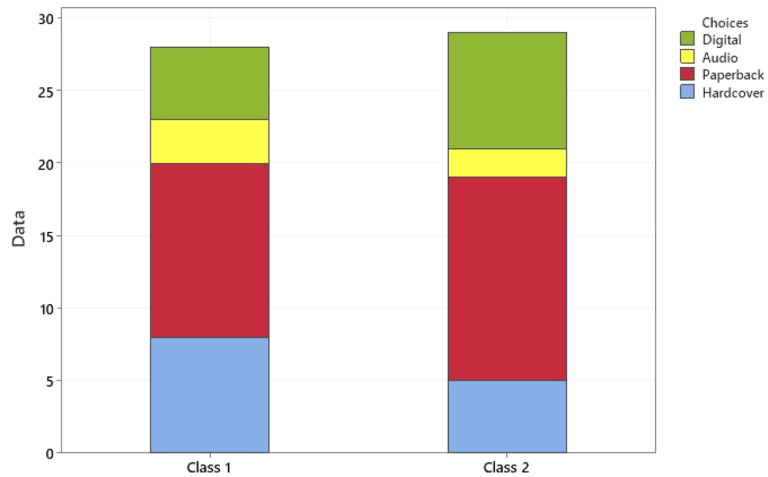
Part D. Given that the contestant was a woman, what is the probability that she won on an episode of Jeopardy!?

Part E. Discuss the content of the article and the results of Part C. and Part D. with a classmate.

Instructional Items

Instructional Item 1

Two different classes are asked their preferred format for reading books. In the first class 8 students said they preferred a hardcover book, 12 said they preferred a paperback book, 3 said they preferred an audio book and 5 said they preferred a digital format. In the second class, 5 students said they preferred a hardcover book, 14 said they preferred a paperback book, 2 said they preferred an audio book and 8 said they preferred a digital format. A segmented bar graph displays this data below.



Part A. What is the probability that a randomly selected student from the two classes prefers a printed format (hardcover or paperback)?

Part B. What is the probability that a randomly selected student is from class 1 and prefers an audio book?

Part C. Given the student is from class 2, what is the probability that he or she prefers a digital format?

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

MA.912.DP.4 *Use and interpret independence and probability.*

MA.912.DP.4.1

Benchmark

MA.912.DP.4.1	Describe events as subsets of a sample space using characteristics, or categories, of the outcomes, or as unions, intersections or complements of other events.

Connecting Benchmarks/Horizontal Alignment

- MA.912.DP.3.1, MA.912.DP.3.2, MA.912.DP.3.3

Terms from the K-12 Glossary

- Conditional relative Frequency
- Event
- Experimental probability
- Frequency table
- Joint frequency
- Sample space
- Theoretical probability

Vertical Alignment

Previous Benchmarks

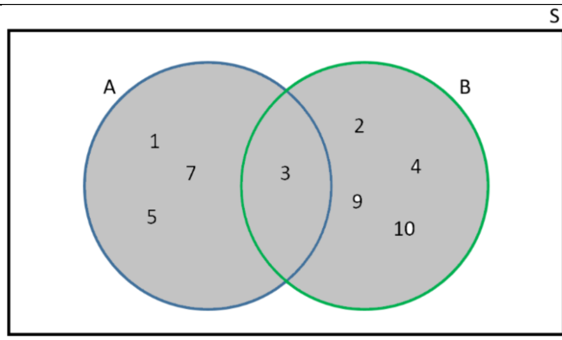
- MA.7.DP.2.1
- MA.8.DP.2.1, MA.8.DP.2.2, MA.8.DP.2.3

Next Benchmarks

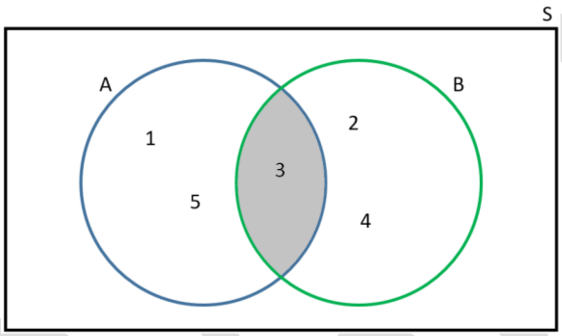
Purpose and Instructional Strategies

In grade 7 and grade 8, students began to identify what made up the sample space and explored basic probability by using theoretical probability for repeated experiments. In Mathematics for College Statistics, students break the sample space up into subsets. Students also identify the union, intersection and complement of other events.

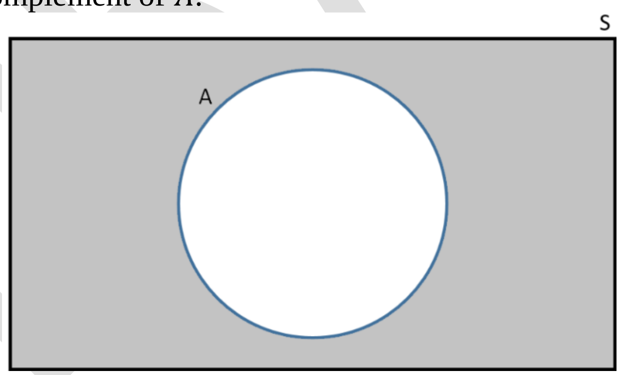
- It is the intention of this benchmark to include basic set notation which corresponds with MA.912.LT.5.4.
- Students may see the complement written with an exponent of C (X^C) or a prime (X').
- It is the intention of this benchmark to include Venn Diagrams, which corresponds with MA.912.LT.5.5.
- It is the intention of this benchmark to include two-way tables.
- It is the intention of this benchmark to include multiple events (2 or more).
- The null/empty set is denoted by { } or $\emptyset$.
- Sets may also be described with a statement.
    - For example, set A consists of prime numbers less than 100.
- Instruction should include the use of diagrams in order to explore and illustrate concepts of complements, unions, intersections, difference and products of two sets.
    - The union of two sets is a set containing all elements that are in A or in B.
        - For example, $A = \{1, 3, 5, 7\}$ and $B = \{2, 3, 4, 9, 10\}$, then $A \cup B$ is $\{1, 3, 5, 7\} \cup \{2, 3, 4, 9, 10\} = \{1, 2, 3, 4, 5, 7, 9, 10\}$. The Venn diagram of this is the following where the shaded region represents $A \cup B$:



- The intersection of two sets is denoted as $A \cap B$ and consists of all elements that are both in A and B.
    - For example, $A = \{1, 3, 5\}$ and $B = \{2, 3, 4\}$, then $A \cap B$ is $\{1, 3, 5\} \cap \{2, 3, 4\} = \{3\}$. The Venn diagram of this is the following where the shaded region represents $A \cap B$:



- The complement of a set can be denoted as A^c, $\bar{A}$, A' or $\sim A$. This is the set of all elements that are in the universal set S but not in A.
    - The Venn diagram below shows that the shaded region represents the complement of A.



Common Misconceptions or Errors

- Students often count the intersection more than once if finding the unions when the events are not mutually exclusive.
- Students may forget about the values that are not in the sets.
- Students may confuse the symbols used in basic set notation.
- Students may have trouble finding the total from a Venn diagram and only count parts of a section, count a section multiple times, or have difficulty understanding the impact a number outside of the circles on Venn diagrams has on the data set.

Instructional Tasks

Instructional Task 1 (MTR.7.1)

The table below provides information on 10 students presenting in the Future Business Leaders of America competition. The students are identified by the order they are presenting in. For each student, their sex, age, whether or not they plan to study business in college, whether or not they play a sport, and how many colleges they plan to apply to.

Presenter Number	Sex	Age	Study Business	Play a Sport	Number of Applications
1	Male	18	Yes	No	3
2	Female	18	Yes	Yes	4
3	Female	17	Yes	No	2
4	Male	18	No	Yes	2
5	Female	18	Yes	No	5
6	Male	17	No	No	7
7	Male	17	Yes	Yes	3
8	Male	18	Yes	Yes	5
9	Female	16	No	Yes	4
10	Male	16	Yes	No	2

Part A. What is the sample space for number of colleges these students plan to apply to?

Part B. What outcomes from the sample space in part A are in the event that the student plans to study business?

Part C. Consider the following 3 events. For each list, which students would be make up the event:

- F = The selected student is female
- S = The selected student plays a sport
- A = The selected student is less than 18 years old

Part D. Based on part C, which outcomes are in the following events?

- $F \cup S$
- $F \cap A$
- A^C

Instructional Items

Instructional Item 1

Given the sets $X = \{3, 5, 7, 9, 11, 13\}$ and $Y = \{1, 2, 3, 4, 5, 9, 10\}$, find the following.

Part A. $X \cup Y$

Part B. $X \cap Y$

Part C. $X \cap Y^C$

Instructional Item 2

The results from a survey about whether students (male or female) at a university were from England, Wales or Scotland is summarized in the following two-way table:

	Male	Female	Total
Scotland	34	16	50
Wales	40	8	48
England	2	0	2
Total	76	24	100

Find the following:

Part A. Male ∪ Female

Part B. Male ∩ England

Part C. England ∩ Male^{C}

The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.

MA.912.DP.4.2

Benchmark

MA.912.DP.4.2	Determine if events A and B are independent by calculating the product of their probabilities.

Connecting Benchmarks/Horizontal Alignment

- MA.912.DP.3

Terms from the K-12 Glossary

- Event

Vertical Alignment

Previous Benchmarks

- MA.7.DP.2.2, MA.7.DP.2.3, MA.7.DP.2.4
- MA.8.DP.2.2, MA.8.DP.2.3

Next Benchmarks

Purpose and Instructional Strategies

In grade 7 and grade 8, students began working with theoretical probabilities and comparing them to experimental probability. In Mathematics for College Statistics, students determine if two events are independent of each other. Independence in this sense means that knowing whether one event occurred does not change the probability of the other event. For this benchmark, students determine independence if the probability of Event A and B is equivalent to the probability of A times the probability of B.

- Instruction includes understanding that events are independent if $P(A \cap B) = P(A) \times P(B)$.
- Instruction includes use conditional probability to check independence (MA.912.DP.4.4). If $P(A|B) = P(A)$ then the events are independent.
- Instruction includes distinguishing independence from mutually exclusive events.
- Instruction identifies that mutually exclusive events $P(A \cap B) = 0$ which means the two events cannot occur at the same time.
- Instruction notes that $P(A \cap B)$ is the same as $P(B \cap A)$.
- Instruction includes the use of technology whether it be a statistical program or graphing calculators.
- Instruction may differentiate by not using set notation to be able to identify if students are struggling with identifying which outcomes are in each event or if they are struggling to interpret set notation.
- Instruction includes the use of real-world data. Although these problem types can still be used, in recent years there has been a push to move away from games of chance when discussing joint probabilities, conditional probabilities and independence in college statistics. The focus is more on contingency tables as this can begin to flow into

hypothesis testing for proportions.

Common Misconceptions or Errors

- Students may confuse what it means to be dependent and independent.
- Students may confuse independence with mutually exclusive events.
- Students may need to recall how to convert fractions, decimals and percentages.

Instructional Tasks

Instructional Task 1 (MTR.6.1, MTR.7.1)

One card is selected at random from a deck of 6 cards. Each card has a number and either a heart or a diamond:{ 3♥, 5♥, 2♦, 9♦, 9♥, 4♦}.

Part A. Let H be the event that the selected card is a heart, and F be the event that the selected card is a 4. Are the events H and F independent? Justify your answer with calculations.

Part B. Let D be the event that the selected card is a diamond, and N be the event that the selected card is a 9. Are the events D and N independent? Justify your answer with calculation.

Instructional Items

Instructional Item 1

The table below shows the response of students to a survey that asked them about the type of pets that they own. Use the table to determine independence.

	Have a Pet Bird	Have a Pet Dog	Have a Pet Cat	Total
Freshman	2	12	7	21
Sophomore	0	16	9	25
Junior	3	15	11	29
Senior	1	14	12	27
Total	6	57	39	102

Part A. Are the events the student is a Freshman and the student has a pet dog independent?

Part B. Are the events the student has a pet bird and the student is a senior independent?

Instructional Item 2

Given the information below, determine whether the events A and B are independent, the events A and C are independent and the events B and C are independent.

$$P(A) = 0.28$$
$$P(B) = 0.54$$
$$P(C) = 0.36$$
$$P(A \cap B) = 0.15$$
$$P(A \cap C) = 0.43$$
$$P(B \cap C) = 0.19$$

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

MA.912.DP.4.3

Benchmark

MA.912.DP.4.3 Calculate the conditional probability of two events and interpret the result in terms of its context.

Connecting Benchmarks/Horizontal Alignment

- MA.912.DP.3

Terms from the K-12 Glossary

- Conditional relative Frequency
- Event
- Experimental probability
- Frequency table
- Joint frequency
- Sample space
- Theoretical probability

Vertical Alignment

Previous Benchmarks

- MA.7.DP.2.1, MA.7.DP.2.2, MA.7.DP.2.3
- MA.8.DP.2.1, MA.8.DP.2.2, MA.8.DP.2.3

Next Benchmarks

Purpose and Instructional Strategies

In grade 7 and grade 8, students began working with theoretical probabilities and comparing them to experimental probability. In Mathematics for College Statistics, students calculate the conditional probability of two events and interpret the results in context of the problem.

- The conditional probability of an event B is the probability that the event will occur given the knowledge that an event A has already occurred. This probability is written $P(B|A)$, notation for the probability of B given A.
  - In the case where events A and B are independent, then $P(B|A) = P(B)$.
  - If events A and B are not independent, then $P(A \text{ and } B) = P(A)P(B|A)$.
- The conditional probability $P(B|A)$ is easily obtained by dividing by $P(A)$:
$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$
- When using the conditional probability formula you would divide by the probability of B if finding the probability of A given B:
$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$
- Conditional probabilities can be observed in tree diagrams as the branch stemming from another branch which may be a helpful visual for some students to understand the meaning of conditional probabilities.
- Instruction includes the use of two-way tables to help determine independence or association (MA.912.DP.3).
- Instruction includes the use of technology whether it be a statistical program or graphing calculators.
- Instruction includes the use of real-world data. While these problem types may still be utilized, in recent years there has been a push to move away from games of chance when discussing joint probabilities, conditional probabilities and independence in college

statistics. The focus is more on contingency tables as this can begin to flow into hypothesis testing for proportions.

Common Misconceptions or Errors

- Students may think the symbol used for conditional probability is a slash that would be used to represent division and simply divide the probability of A by the probability of B.
- Students may get confused as to which event probability should be the denominator.
- Students may get confused when working with a two-way table that they need to restrict their answer to a certain section that is from the "given" conditional piece.
    - For example, when given the condition of male they are only looking in the row or column containing males to get the total.
- Students who have difficulty with the terminology and notation will also have difficulty in understanding what is being asked by the questions.

Instructional Tasks

Instructional Task 1 (MTR.7.1)

On April 15, 1912, the Titanic struck an iceberg and rapidly sank with only 710 of her 2,204 passengers and crew surviving. Data on survival of passengers are summarized in the table below. (Source: http://www.encyclopedia-titanica.org/titanic-statistics.html)

	Survived	Did not Survive	Total
First class passengers	201	123	324
Second class passengers	118	166	284
Third class passengers	181	528	709
Total passengers	500	817	1317

Calculate the following probabilities. Round your answers to three decimal places.

Part A. If one of the passengers is randomly selected, what is the probability that this passenger was in first class?

Part B. If one of the passengers is randomly selected, what is the probability that this passenger survived?

Part C. If one of the passengers is randomly selected, what is the probability that this passenger was in first class and survived?

Part D. If one of the passengers is randomly selected, what is the probability that the passenger survived, given that this passenger was in first class?

Part E. If one of the passengers is randomly selected, what is the probability that the passenger was in first class, given that this passenger survived?

Part F. If one of the passengers who survived is randomly selected, what is the probability that this passenger was in third class?

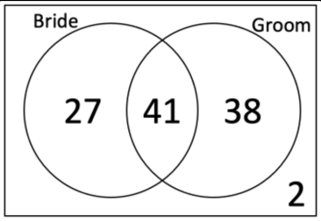
Part G. Why is the answer to Part D larger than the answer to Part C?

Part H. Why is the answer to Part D different from the answer in Part E?

Instructional Items

Instructional Item 1

The usher at a wedding asks each guest if they are a friend of the bride or the groom. The results are given in the Venn Diagram below.



Given that a randomly selected guest is a friend of the bride, find the probability that they are a friend of the groom.

Instructional Item 2

Andre surveyed students at his school and found the following probabilities:

$$P(Senior) = \ 0.23$$
$$P(Taken\ ACT) = 0.64$$
$$P(Taken\ SAT) = 0.79$$
$$P(Senior\ and\ ACT) = 0.53$$
$$P(Senior\ and\ SAT) = 0.61$$

Part A. Find the probability that the student has taken the ACT given they are a Senior.

Part B. Find the probability that the student is a Senior given they have taken the SAT.

Instructional Item 3

The results from a survey about whether students (male or female) at a university were from England, Wales or Scotland is summarized in the following two-way table.

	Male	**Female**	**Total**
Scotland	34	16	50
Wales	40	8	48
England	2	0	2
Total	76	24	100

Find the following:

Part A. $P(Wales|Male)$

Part B. $P(Female|Scotland)$

Part C. $P(Male|Scotland)$

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

MA.912.DP.4.4

Benchmark

MA.912.DP.4.4 Interpret the independence of two events using conditional probability.

Connecting Benchmarks/Horizontal Alignment

- MA.912.DP.3

Terms from the K-12 Glossary

- Conditional relative Frequency
- Event
- Experimental probability
- Frequency table
- Joint frequency
- Sample space
- Theoretical probability

Vertical Alignment

Previous Benchmarks

- MA.7.DP.2.2, MA.7.DP.2.3, MA.7.DP.2.4
- MA.8.DP.2.2, MA.8.DP.2.3

Next Benchmarks

Purpose and Instructional Strategies

In grade 7 and grade 8, students began working with theoretical probabilities and comparing them to experimental probability. In Mathematics for College Statistics, students determine if two events are independent of each other. Independence in this sense means that knowing whether one event occurred does not change the probability of the other event. For this benchmark, students determine independence using conditional probabilities. Two events, A and B, are independent if $P(A \mid B) = P(A)$ and $P(B \mid A) = P(B)$.

- Instructions includes using the product of probabilities to check independence (MA.912.DP.4.2). If $P(A \cap B) = P(A) \times P(B)$, then the events are independent.
- Instruction distinguishes independence from mutually exclusive events in order to differentiate different types of joint probabilities.
- Instruction clarifies that mutually exclusive events $(A \cap B) = 0$, which means the two events cannot occur at the same time.
- Instruction includes examples that demonstrate that $P(A \mid B) \neq P(B \mid A)$ unless $P(A) = P(B)$.
- Instruction notes that when we check for independence in real world data sets, it is rare to get perfectly equal probabilities. We often assume that events are independent and test that assumption on sample data. If the probabilities are significantly different, then we conclude the events are not independent.
- Instruction includes the use of technology whether it be a statistical program or graphing calculators.
- Instruction includes the use of real-world data. In recent years there has been a push to move away from games of chance when discussing joint probabilities, conditional probabilities and independence in college statistics. The focus is more on contingency tables as this can begin to flow into hypothesis testing for proportions.

Common Misconceptions or Errors

- Students may confuse what it means to be dependent and independent.
- Students may confuse independence with mutually exclusive events.
- Students may have difficulty recalling how to convert fractions, decimals and percentages.
- Students may think the symbol used for conditional probability is a slash that would be used to represent division and simply divide the probability of A by the probability of B.
- Students may get confused as to which event probability should be the denominator.
- Students may get confused when working with a two-way table that they need to restrict their answer to a certain section that is from the "given" conditional piece.
    - For example, when given the condition of male they are only looking in the row or column containing males to get the total.
- Students who have difficulty with the terminology and notation will also have difficulty in understanding what is being asked by the questions.

Instructional Tasks

Instructional Task 1

On April 15, 1912, the Titanic struck an iceberg and rapidly sank with only 710 of her 2,204 passengers and crew surviving. Data on survival of passengers are summarized in the table below. (Source: http://www.encyclopedia-titanica.org/titanic-statistics.html)

	Survived	Did not Survive	Total
First class passengers	201	123	324
Second class passengers	118	166	284
Third class passengers	181	528	709
Total passengers	500	817	1317

Part A. Are the events "passenger survived" and the "passenger was in first class" independent events?

Part B. Are the events "passenger survived" and the "passenger was in third class" independent events?

Part C. Did all passengers aboard the Titanic have the same probability of surviving? Defend your response using calculations. Be prepared to share your reasoning.

Instructional Items

Instructional Item 1

The hiring department of Trinity United Mortgage Company compiled the accompanying data regarding the income and education of its employees.

	Income Below \$60,000	Income Over \$60,000
Noncollege Graduate	5	8
College Graduate	11	17

Using conditional probabilities, are the events "employee is a college graduate" and the "employee makes over \$60,000" independent? Explain.

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

MA.912.DP.4.5

Benchmark

MA.912.DP.4.5	Given a two-way table containing data from a population, interpret the joint and marginal relative frequencies as empirical probabilities and the conditional relative frequencies as empirical conditional probabilities. Use those probabilities to determine whether characteristics in the population are approximately independent.

Example: A company has a commercial for their new grill. A population of people are surveyed to determine whether or not they have seen the commercial and whether or not they have purchased the product. Using this data, calculate the empirical conditional probabilities that a person who has seen the commercial did or did not purchase the grill.

Benchmark Clarifications:
Clarification 1: Instruction includes the connection between mathematical probability and applied statistics.

Connecting Benchmarks/Horizontal Alignment

- MA.912.FL.1.1
- MA.912.DP.1.2
- MA.912.DP.3

Terms from the K-12 Glossary

- Categorical data
- Bivariate data
- Conditional relative Frequency
- Event
- Experimental probability
- Joint relative frequency

Vertical Alignment

Previous Benchmarks

- MA.7.DP.1.3
- MA.7.DP.2.4
- MA.8.DP.2.1, MA.8.DP.2.3

Next Benchmarks

Purpose and Instructional Strategies

In grade 7 and grade 8, students were introduced to collecting data from repeated experiments and organizing them into tables as well as making predictions based on proportional relationships. In Algebra I, students studied bivariate categorical data and displayed it in tables showing joint frequencies and marginal frequencies. In Mathematics for College Statistics, students continue this exploration with the use of real-world data and focusing on conditional relative frequencies.

- Instruction includes the connection to MA.912.DP.3.1, where students construct two-way frequency tables to summarize bivariate categorical data.
- Empirical probability is an estimate of a probability that comes from the ratio of the number of outcomes in which a specified event occurs to the total number of trials, and is also referred to as the experimental probability.
- In later courses, students may hear the term for two-way tables as contingency tables. Instruction should include the use of both terms interchangeably so that students are familiar with them.

- Students are expected to be able to apply the formulas for checking if two events are independent in order to determine whether characteristics in the population are approximately independent.
    - If $P(A \mid B) = P(A)$, then events A and B are independent.
    - If $P(A \cap B) = P(A) * P(B)$, then the events A and B are independent.
- When we check for independence in real-world data sets, it is rare to get perfectly equal probabilities. We often assume that events are independent and test that assumption on sample data. If the probabilities are significantly different, then we conclude the events are not independent.
- Students are expected to interpret the joint and marginal frequencies and therefore must know the differences between these two.
    - **Marginal frequencies**
      Entries in the "Total" rows and "Total" columns are called marginal frequencies. The areas highlighted in yellow in the table below represent the marginal frequencies.

	7	5	
	2	7	

    - **Joint frequencies**
      Entries in the body of the table represent joint frequencies. The area highlighted in yellow in the table represents the joint frequencies.

			12
			9
	9	12	21

Common Misconceptions or Errors

- Students may confuse joint relative frequencies, marginal relative frequencies and conditional relative frequencies.
- Students may leave off the total row or total column.
- Students may have difficulty when relative frequencies are given in decimal form.
- When making conditional relative frequency tables, students may choose the wrong value to be the denominator when converting to relative frequencies.
- Students may need remediation to convert fractions, decimals and/or percentages.
- Student forget the formulas they can use to check for independence.

Instructional Tasks

Instructional Task 1 (MTR.2.1, MTR.7.1)

Students in Mr. Williams class answered a survey question on the type of movie they prefer more between comedy and action. The data are shown in the table below.

	Comedy	**Action**	**Total**
Male	8	12	20
Female	8	4	12

Total	16	16	32

Part A. What could be said about the marginal relative frequencies in the total row?
Part B. What could be said about the marginal relative frequencies in the total column?
Part C. What could we conclude if we look at the joint relative frequencies?
Part D. For what type of movie would you say that the conditional relative frequencies for females and males are very different? Explain.
Part E. For what type of movie would you say that the conditional relative frequencies for females and males are nearly equal? Explain.
Part F. If you were a movie producer, what type of movie would you produce based on these results and why.

Instructional Task 2 (MTR.4.1, MTR.7.1)

On April 15, 1912, the Titanic struck an iceberg and rapidly sank with only 710 of her 2,204 passengers and crew surviving. Some believe that the rescue procedures favored the wealthier first class passengers. Other believe that the survival rates can be explained by the “women and children first” policy. Data on survival of passengers are summarized in the table below. Based on the data, can you find evidence to support the procedures favored the wealthier passengers? Based on the data, can you find evidence to support the procedures favored women and children? Explain. (Source: [Titanic Statistics])

	Survived	**Did not Survive**	**Total**
Children in first class	4	1	5
Women in first class	139	4	143
Men in first class	58	118	176
Children in second class	22	0	22
Women in second class	83	12	95
Men in second class	13	154	167
Children in third class	30	50	80
Women in third class	91	88	179
Men in third class	60	390	450
Total passengers	500	817	1317

Instructional Items

Instructional Item 1

A company is running a discount via ads on social media. A population of people are surveyed to determine whether or not they have seen the ad and whether or not they have purchased any products. Using this data, calculate the empirical conditional probabilities that a person who has seen the ad did or did not purchase a product.

	Saw the Ad	**Did not See the Ad**
Purchased at least one item	42	9
Didn’t purchase any items	19	30

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

MA.912.DP.4.6

Benchmark

MA.912.DP.4.6 Recognize and explain the concepts of conditional probability and independence in everyday language and everyday situations.

Connecting Benchmarks/Horizontal Alignment

- MA.912.DP.4.2, MA.912.DP.4.4, MA.912.DP.4.5, MA.912.DP.4.8

Terms from the K-12 Glossary

- Event
- Conditional relative frequency

Vertical Alignment

Previous Benchmarks

- MA.7.DP.2.2, MA.7.DP.2.3, MA.7.DP.2.4
- MA.8.DP.2.2, MA.8.DP.2.3

Next Benchmarks

Purpose and Instructional Strategies

In grade 7 and grade 8, students began working with theoretical probabilities and comparing them to experimental probability. In Mathematics for College Statistics, students determine if two events are independent of each other. Independence in this sense means that knowing whether one event occurred does not change the probability of the other event. For this benchmark, students apply what they have learned previously about conditional probability and independence to explain what they mean in context of the situation using everyday language.

- Instructions includes using the product of probabilities to check independence (MA.912.DP.4.2). If $P(A \cap B) = P(A) \times P(B)$, then the events are independent.
- Two events, A and B, are independent if $(A \mid B) = P(A)$ and $P(B \mid A) = P(B)$.
- Students are expected to be able to apply the formulas for checking if two events are independent in order to determine whether characteristics in the population are approximately independent.
- Students should either memorize or be able to derive the rules used for checking for independence.
- Instruction distinguishes independence from mutually exclusive events in order to differentiate different types of joint probabilities.
- Students should understand that $(A \mid B) \neq P(B \mid A)$ unless $P(A) = P(B)$.
- Be sure to distinguish independence from mutually exclusive events.
- In mutually exclusive events $(A \cap B) = 0$, which means the two events cannot occur at the same time.
- Students should understand that $P(A \cap B)$ is the same as $P(B \cap A)$.
- Conditional probability is a measure of the probability of an event occurring, given that another event has already occurred and can be found using the formula $P(A|B) = \frac{P(A \cap B)}{P(B)}$.
- When we check for independence in real-world data sets, it is rare to get perfectly equal probabilities. We often assume that events are independent and test that assumption on sample data. If the probabilities are significantly different, then we conclude the events are not independent.
- Instruction includes the use of technology whether it be a statistical program or graphing calculators.
- Instruction includes the use of real-world data. In recent years, there has been a push to move away from games of chance when discussing joint probabilities, conditional

probabilities, and independence in college statistics. The focus is more on contingency tables as this can begin to flow into hypothesis testing for proportions.

Common Misconceptions or Errors

- Students may confuse what it means to be dependent and independent.
- Students may confuse independence with mutually exclusive events.
- Students may need remediation to convert fractions, decimals and percentages.
- Student may mix up which event is used as the denominator in conditional probability problems.
- Students may have difficulty interpreting independence and conditional probability when events use negative adverbs like not or did not.

Instructional Tasks

Instructional Task 1 (MTR.3.1, MTR.7.1)

Before tests in economics class, Chase sometimes studies and sometimes does not. After taking a course on probability, Chase says, “The events ‘I study for the economics test’ and ‘I get a high score on the test’ are independent.” Explain what this means in terms of the relationship between Chase studying for economics and his probability of getting a high score on the test in language that someone who hasn’t taken statistics would understand. After making this statement, Chase kept track of whether or not he studied for the economics test and his scores on the tests for the rest of the school year. He found that on the test he studied for, he got a high score on the test 78% of the time, and on the test he didn’t study for, he got a high score on the test 25% of the time. Do these finding indicate that the events, studied for the economics test and got a high score on the test, are independent?

Instructional Task 2 (MTR.7.1)

In the game of basketball, it is often wondered whether two shots are independent or dependent: does the probability of making the second free throw depend on whether a player makes the first free throw? After analyzing the data for a specific basketball player in the WNBA, statisticians determined that her first and second free throws are entirely independent events. The frequency table below shows the data that analysts used to determine this independence.

Part A. Fill in the missing values from the frequency table.

	Makes First Shot	Misses First Shot	Total
Makes Second Shot			154
Misses Second Shot			66
Total	190	30	220

Part B. Explain what it means that her first and second free throws are entirely independent events.

Instructional Items

Instructional Item 1

In the game of basketball, it is often wondered whether two shots are independent or dependent: does the probability of making the second free throw depend on whether a player makes the first free throw? The frequency table below shows the data for a specific basketball player in the WNBA.

	Makes First Shot	Misses First Shot	Total

Makes Second Shot	133	21	154
Misses Second Shot	57	9	66
Total	190	30	220

Fill in the missing values from the frequency table. What evidence can you use to show the events are independent?

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

MA.912.DP.4.7

Benchmark

MA.912.DP.4.7	Apply the addition rule for probability, taking into consideration whether the events are mutually exclusive, and interpret the result in terms of the model and its context.

Connecting Benchmarks/Horizontal Alignment

- MA.912.DP.3.1, MA.912.DP.3.2, MA.912.DP.3.5

Terms from the K-12 Glossary

- Conditional relative frequency
- Event
- Experimental probability
- Frequency table
- Joint frequency
- Joint relative frequency
- Sample space
- Theoretical probability

Vertical Alignment

Previous Benchmarks

- MA.7.DP.2.2, MA.7.DP.2.3, MA.7.DP.2.4
- MA.8.DP.2.2, MA.8.DP.2.3

Next Benchmarks

Purpose and Instructional Strategies

In grade 7 and grade 8, students began working with theoretical probabilities and comparing them to experimental probability as well as collecting data through simulation. In Mathematics for College Statistics, students apply the addition rule for probability, $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$, to calculate the probability of event A or B. Students recognize the effect mutually exclusive events have on this rule and be able to interpret the result in terms of the model and its context.

- Students should be able to calculate the union of two events using the formula $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
- The symbol for union, ∪, is read as "or" and the symbol for intersection, ∩, is read as "and." If students are trying to identify what action to take from a word problem involving probability, "or" should be the key word letting students know to use the addition rule.

- Instruction includes the use of a reference sheet with the formulas for the addition rule for probability, the multiplication rule for probability and the conditional probability formula.
- Students should be able to explain why we need to subtract the intersection when calculating the union.
- Be sure to distinguish independence from mutually exclusive events.
- In mutually exclusive events, $P(A \cap B) = 0$, which means the two events cannot occur at the same time.
- In mutually exclusive events, $P(A \cup B) = P(A) + P(B)$.
- Instructions includes using the product of probabilities to check independence (MA.912.DP.4.2). If $P(A \cap B) = P(A) \times P(B)$, then the events are independent.
- Instruction distinguishes independence from mutually exclusive events in order to differentiate different types of joint probabilities.
- Instruction clarifies that mutually exclusive events $(A \cap B) = 0$, which means the two events cannot occur at the same time.
- Instruction includes the use of models, which may include Venn Diagrams, organized lists, two-way tables or tree diagrams.
- Students are expected to be able to interpret the results in context as well as how it would be represented on a model such as a Venn Diagram.
- In a college statistics class, students will typically work with the addition rule for probability and mutually exclusive events using contingency tables and Venn diagrams.
- Students should be able to work backwards if given the union to find other missing pieces in the addition rule formula.
- Instruction includes the use of technology whether it be a statistical program or graphing calculators.
- Instruction includes the use of real-world data.

Common Misconceptions or Errors

- Students may not recognize events as mutually exclusive or non-mutually exclusive.
- Student may have difficulty when pulling information from a Venn Diagram, especially when they have multiple sections that overlap.
- Students may make errors with determining the correct denominator. They may use the total rather than the specified event.
- Students may have misconceptions in their understanding of the "overlap" in compound events.
- Students may have difficulty with algebraic skills if using the addition rule formula to work backwards and solve for an unknown.

Instructional Tasks

Instructional Task 1 (MTR.7.1)

In a certain school, 15 percent of the students are enrolled in a physics course, 27 percent are enrolled in a foreign language course, and 30 percent are enrolled in either a physics course or a foreign language course or both.

Part A. What is the probability that a student chosen at random from this school will be enrolled in a physics course?

Part B. What is the probability that a student chosen at random from this school will be enrolled in a foreign language course?

Part C. What is the probability that a student chosen at random from this school will be enrolled in both a foreign language course and a physics course?

Part D. Draw a Venn Diagram to model this situation

Instructional Task 2 (MTR.4.1, MTR.7.1)

If a fair single 6-sided die is rolled, answer the following.

Part A. What is the probability of rolling a number less than 4 or an even number?

Part B. What is the probability of rolling a number greater than 5 or an odd?

Part C. Using statistical terminology, explain why you needed to subtract in Part A. but not for Part B.

Instructional Items

Instructional Item 1

At The Triple Stack cafe, everyone orders either pancakes, waffles, or French toast. Let $S =$ the event that a randomly selected customer puts syrup on their food. Let $F =$ the event that a randomly selected customer orders a fruit topping on their food. Suppose that after years of collecting data, The Triple Stack cafe has estimated the following probabilities:

$$P(S) = 0.8$$
$$P(F) = 0.5$$
$$P(S \text{ or } F) = 0.9$$

Estimate $P(S \text{ and } F)$ and interpret this value in the context of the problem.

Instructional Item 2

Given the two-way table below, find the following probabilities and draw a Venn-diagram to model each situation.

	Ride the Bus	Do Not Ride the Bus	Total
Boys	7	5	12
Girls	2	7	9
Total	9	12	21

Part A. Find the probability that a randomly selected person rides the bus or is a girl.

Part B. Find the probability that a randomly selected person is a boy or does not ride the bus.

Part C. What are two events from the table that are not mutually exclusive?

Part D. Why are the events 'Ride the Bus' and 'Do Not Ride the Bus' mutually exclusive?

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

MA.912.DP.4.8

Benchmark

MA.912.DP.4.8	Apply the general multiplication rule for probability, taking into consideration whether the events are independent, and interpret the result in terms of the context.

Connecting Benchmarks/Horizontal Alignment

- MA.912.DP.3.1, MA.912.DP.3.2, MA.912.DP.3.5

Terms from the K-12 Glossary

- Conditional relative Frequency

- Event
- Experimental probability
- Frequency table
- Joint frequency
- Joint relative frequency
- Sample space
- Theoretical probability

Vertical Alignment

Previous Benchmarks
- MA.7.DP.2.2, MA.7.DP.2.3, MA.7.DP.2.4
- MA.8.DP.2.2, MA.8.DP.2.3

Next Benchmarks

Purpose and Instructional Strategies

In grades 7 and 8, students began working with theoretical probabilities and comparing them to experimental probability as well as collecting data through simulation. In Mathematics for College Statistics, students have previously learned how to determine if events are independent, calculate conditional probabilities and apply the addition rule for probability.

- Students apply the general multiplication rule for probability, $P(A \text{ } and \text{ } B) = P(A) * P(B \text{ } given \text{ } A)$, to calculate the probability of the intersection of events A and B. Students recognize the effect independent events have on this rule and be able to interpret the result in terms of its context.
- Students should be able to calculate the intersection of two events using the formula $P(A \cap B) = P(A) * P(B|A)$.
- Students should be able to explain that the intersection is referring to the events occurring simultaneously.
- Be sure to distinguish independence from mutually exclusive events.
- In mutually exclusive events, $P(A \cap B) = 0$, which means the two events cannot occur at the same time.
- In independent events, $P(A|B) = P(A)$ and $P(B|A) = P(B)$, which means the occurrence of one event doesn't change the probability of the occurrence of the other.
- If events are independent, it simplifies the multiplication rule to $P(A \cap B) = P(A) * P(B)$.
- Instructions includes using the product of probabilities to check independence (MA.912.DP.4.2). If $P(A \cap B) = P(A) \times P(B)$, then the events are independent.
- Instruction distinguishes independence from mutually exclusive events in order to differentiate different types of joint probabilities.
- Instruction clarifies that mutually exclusive events $(A \cap B) = 0$, which means the two events cannot occur at the same time.
- Instruction includes the use of models, which may include Venn Diagrams, organized lists, two-way tables or tree diagrams.
- Students should be able to work backwards if given the intersection to find other missing pieces in the general multiplication rule formula.
- Students are not expected to memorize the formulas for the addition rule for probability, the multiplication rule for probability or the conditional probability formula.

- Instruction includes the use of technology whether it be a statistical program or graphing calculators.
- Instruction includes the use of real-world data.

Common Misconceptions or Errors

- Students may add the probabilities when they should be multiplying.
- Students may assume the events are independent and use the wrong formula.
- Students may confuse the union (∪) and intersection (∩) symbols.
- Students may mix up mutually exclusive events with independent events.
- Student may have difficulty when pulling information from a Venn Diagram especially when they have multiple sections that overlap.
- Students may make errors with determining the correct denominator. They may use the total rather than the specified event.
- Students may have difficulty with algebraic skills if using the addition rule formula to work backwards and solve for an unknown.
- Students may use the wrong probabilities if they are asked to find the probability of neither where they need to find the complements of the events occurring simultaneously.

Instructional Tasks

Instructional Task 1 (MTR.7.1)

Suppose you have a bag with 6 green marbles, 3 pink marbles and 5 yellow marbles. Answer the following questions and clearly show your process.

Part A. What is the probability of pulling out a pink marble followed by a green marble if you are going to pull out one marble, record its color, put it back in the bag and draw another marble.

Part B. What is the probability of pulling out a pink marble followed by a green marble if you are going to pull out one marble, record its color, leave it out and draw another marble.

Part C. Why are these probabilities different when you are first pulling a pink and then a green in both cases?

Part D. What is the probability of pulling out two marbles without replacement and neither one being yellow?

Instructional Task 2 (MTR.2.1, MTR.4.1, MTR.7.1)

Given the data collected from elementary students at Trinity Oaks Elementary School, interpret as much as you can about the highlighted cell.

	Ride the Bus	**Do Not Ride the Bus**	**Total**
Boys	7	5	12
Girls	2	7	9
Total	9	12	21

Instructional Items

Instructional Item 1

A certain brand of at-home gluten sensitivity test has a 5% chance of giving a false positive; that is, if they are negative for gluten sensitivity, there is a 5% chance the test will incorrectly give a positive reading. If a person is positive for gluten sensitivity, there is a 1% chance the test will incorrectly give a negative reading (false negative). Suppose the test is given to a group where 2% of the people have a confirmed case of gluten sensitivity. What is the probability that a person in the group is negative for gluten sensitivity, but has a positive test result?

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

MA.912.DP.4.9

Benchmark

MA.912.DP.4.9 Apply the addition and multiplication rules for counting to solve mathematical and real-world problems, including problems involving probability.

Connecting Benchmarks/Horizontal Alignment

Terms from the K-12 Glossary

- Event
- Experimental probability
- Sample space
- Theoretical probability

Vertical Alignment

Previous Benchmarks

- MA.7.DP.2
- MA.8.DP.2

Next Benchmarks

Purpose and Instructional Strategies

In grades 7 and 8, students began working with theoretical probabilities and comparing them to experimental probability as well as collecting data through simulation. In Mathematics for College Statistics, students are introduced to combinatorics by using the addition principle and multiplication principle to figure out how many possibilities they could have in certain situations.

- The multiplication rule or multiplication principle is also referred to as the fundamental counting principle.
- According to the Addition Principle, if one event can occur in m ways and a second event with no common outcomes can occur in n ways, then the first *or* second event can occur in $m + n$ ways.
- According to the Multiplication Principle, if one event can occur in m ways and a second event can occur in n ways after the first event has occurred, then the two events can occur in $m \times n$ ways.
- Instruction includes the use of models, which may include organized lists, two-way tables or tree diagrams.
- Students are not expected to memorize the formulas for the addition rule for probability, the multiplication rule for probability and the conditional probability formula.
- Instruction includes the use of technology whether it be a statistical program or graphing calculators.
- Instruction includes the use of real-world situations and general mathematical problems with no context.

Common Misconceptions or Errors

- Students may initially be unable to identify when they should be using either of these principles.
- Students may get questions incorrect due to overthinking because of the simplicity of the rules.

Instructional Tasks

Instructional Task 1 (MTR.7.1)

Hazel has created a capsule wardrobe made up of shirts, pants, and shoes that can be mixed and matched to make several outfits. She has the following items in her wardrobe:

- 5 different shirts – black stripe, red, yellow, pink, floral print
- 3 pairs of pants – blue jeans, black, grey
- 3 different shoes – sandals, wedges, high heels

Draw a picture or create a tree diagram and use these to determine:

Part A. The number of possible outfits Hazel can create
Part B. The probability of Hazel wearing a red shirt with grey pants and sandals
Part C. The probability of Hazel wearing a red shirt
Part D. The probability of Hazel wearing a floral shirt with blue jeans
Part E. The probability of Hazel wearing wedges
Part F. The probability of Hazel wearing a yellow shirt with black pants

Instructional Items

Instructional Item 1

Suppose you roll a 6-sided die and draw a card from a deck of 52 cards. How many total outcomes are there?

Instructional Item 2

Suppose you roll two 6-sided dice and the outcome is the sum of the two dice. Make a list of all the possible outcomes. How many total outcomes exist?

Instructional Item 3

The menu for a wedding reception has 2 vegetarian meal options and 3 meat meal options. What is the total number of meal options?

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

MA.912.DP.4.10

Benchmark

MA.912.DP.4.10 Given a mathematical or real-world situation, calculate the appropriate permutation or combination.

Connecting Benchmarks/Horizontal Alignment

Terms from the K-12 Glossary

- Conditional relative Frequency
- event
- Experimental probability

- Sample space
- Theoretical probability

Vertical Alignment

Previous Benchmarks

- MA.7.DP.2
- MA.8.DP.2

Next Benchmarks

Purpose and Instructional Strategies

In grades 7 and 8, students determined the sample space and began working with theoretical probabilities and comparing them to experimental probability as well as collecting data through simulation. In Mathematics for College Statistics, students are introduced to combinatorics by using the addition principle and multiplication principle to figure out how many possibilities they could have in certain situations. They will use those principles to calculate permutations and combination.

- In permutation the order of items is important. Rearrangements of the same items are considered to be different sequences.
- Permutations that allow for repetition use the formula n^r, and if repetitions are not allowed, use the formula $\frac{n!}{(n-r)!}$.
  - Different notation for permutations include $P(n, r) = nPr$.
- In combinations the items do not have to be in any particular order. Rearrangements of the same items are considered to be the same sequences.
- Combinations that allow for repetition use the formula $\frac{(r+n-1)!}{r!(n-1)!}$, and if repetitions are not allowed, use the formula $\frac{n!}{r!(n-r)!}$.
  - Different notation for combination include $C(n, r) = nCr = \binom{n}{r}$.
- In the formulas, n is the number of items to choose from and r is how many items you choose.
- Instruction focuses primarily on situations that do not allow repetition with combinations.
- The multiplication principle is also referred to as the fundamental counting principle.
- Instruction includes the use of models, which may include organized lists, two-way tables or tree diagrams.
- Students are not expected to memorize the formulas for permutations and combinations should be given
- Instruction includes the use of technology whether it be a statistical program or graphing calculators.
- Instruction includes the use of real-world situations and general mathematical problems with no context.

Common Misconceptions or Errors

- Students may initially have difficulty in determining if they should be using the formula for combination or permutation.
- Students may make errors when making calculations or using their graphing utility to solve these problems.

Instructional Tasks

Instructional Task 1

An ice cream shop has 12 unique ice cream flavors, 3 different cake flavors and 8 different types of toppings. They offer sundaes and ice cream cake stackers. For the sundaes, the order of the flavors and toppings does not matter but the order is extremely important for the cake stackers. Use this information to find the number of possibilities for each menu item described below.

Menu Item	Description	Number of Possibilities
The Single	One flavor of ice cream with one topping	
Two Tops	One flavor of ice cream with two toppings	
Triple Triple	Three flavors of ice cream with three toppings	
Don't Hold Back	Up to three flavors of ice cream and as many toppings as you want	
The Simple Cake Stack	One cake flavor and one ice cream flavor	
The Stack Stack	A bottom cake flavor a bottom ice cream flavor a top cake flavor and a top cake ice cream	
The Empire	All three cake flavors are used in your choice of order and three ice cream flavors in your choice of order	

Instructional Items

Instructional Item 1

Eight swimmers are competing in the final round of the Olympic 200 meter competition. In how many ways can 3 of the swimmers finish first, second and third to win the gold, silver and bronze medals?

Instructional Item 2

How many different 2-letter arrangements can be selected from the 5 letters in the word PLANT?

Instructional Item 3

Calculate ${}_{12}C_4$ and ${}_{8}P_5$.

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

MA.912.DP.5 *Determine methods of data collection and make inferences from collected data.*

MA.912.DP.5.1

Benchmark

MA.912.DP.5.1 Distinguish between a population parameter and a sample statistic.

Connecting Benchmarks/Horizontal Alignment

- MA.912.DP.1.2

Terms from the K-12 Glossary

- Data
- Population (in data analysis)
- Random sampling
- Statistical question

Vertical Alignment

Previous Benchmarks

- MA.6.DP.1.1
- MA.7.DP.1.3

Next Benchmarks

Purpose and Instructional Strategies

In grades 6 and 7, students examined creating statistical questions to formulate data, and they looked into different calculations that can produced from data. In Mathematics for College Statistics, students explore the differences in populations/samples and parameters/statistics.

- Instruction defines a population to be all of a particular group, while a sample is a subset of the population.
  - For example, a population could be all students who attend Sunset High School, while a sample could be 100 randomly sampled students.
- Instruction emphasizes that calculations from data from an entire population produce parameters. Calculations from data derived from a sample produce statistics.
  - If 29.6% of all students at Sunset High School are freshman, 29.6% would be a parameter. If 32% of our random sample are freshman, then 32% would be a statistic.
- Instruction includes the idea that parameters and statistics can include percentages and modes (for categorical data) and means, medians, standard deviations, interquartile ranges, quartiles and ranges (for numerical data).
- Instruction includes examples where students identify populations, samples, parameters, and statistics.
  - For example, suppose that recent election results show that 49% of all eligible Leon County residents actually voted in the most recent election. A random sample of 200 randomly sample eligible Leon County residents is taken, and 53% say they intend to vote in the next election. The population would be all eligible Leon County residents, with a population parameter of 49%. The sample would be the 200 eligible Leon county residents, with 53% as the sample statistic.
- Instruction notes the difficulty in collecting data from every individual in a population, which therefore, makes actually calculating a true population parameter very difficult. Since samples are easier to obtain, statistics can be calculated more easily and then used to estimate a population parameter. Statistics are often referred to as estimators because of this.
- Instruction includes a discussion about the natural variation that occurs when comparing statistics to parameters and when comparing the statistics of one sample to the statistics of a different sample.

Common Misconceptions or Errors

- Students often have difficulty grasping that statistics and parameters are actual calculations. Students will sometimes confuse the parameter with the actual population or confuse the statistic with the actual sample. Other times they will misinterpret statistics and parameters as data instead of the calculations that result from the data.

Instructional Tasks

Instructional Task 1 (MTR.4.1)

Part A. Suppose we want to know the mean commute time of all employees who live in Florida. Identify the population of interest. Do you think it is possible to collect data from all employees who live in Florida? Explain your reasoning.

Part B. Using the scenario above, what is the population parameter of interest? Do you think it is possible to actually calculate this particular value? Explain why or why not.

Part C. Suppose we randomly sample 2000 employees from all over the state of Florida, and we calculate the mean commute time of these 2000 residents to be 26.8 minutes. Identify the sample and the statistic.

Part D. Do you think the statistic from Part C would be an exact match for the population parameter we are interested in? Explain your reasoning.

Instructional Items

Instructional Item 1

An elementary school principal wants to know the typical amount of time that the parents of students at her school spend reading to their children each week. She sends home a parent survey with each child, and of the 173 surveys returned, she calculates the median to be 50 minutes of reading each week.

Part A. Identify the population and parameter of interest.

Part B. Identify the sample and statistic of interest.

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

MA.912.DP.5.2

Benchmark

MA.912.DP.5.2	Explain how random sampling produces data that is representative of a population.

Connecting Benchmarks/Horizontal Alignment

Terms from the K-12 Glossary

- Data
- Population (in data analysis)
- Random sampling
- Statistical question

Vertical Alignment

Previous Benchmarks

- MA.6.DP.1.1
- MA.7.DP.1.3

Next Benchmarks

Purpose and Instructional Strategies

In grades 6 and 7, students examined formulating statistical questions to formulate data and used proportional relationships to make predictions about a population. In Mathematics for College Statistics, students explore how they can extend this idea to random sampling in order to produce data that is representative of the population.

- Instruction references MA.912.DP.5.1 to reinforce that often times a population is too large to collect data from each individual; a sample is selected as a subset of the population. In order to produce a sample that has similar demographics to the population, a random sample should be selected.
- Instruction explains how convenience samples and voluntary response samples lead to biased samples that most likely do not represent the entire population.
    - For example, if a college professor talks to the first 50 students entering the on-campus library on a Monday morning, this would be a biased convenience sample that would not share the characteristics of all students in the population. Students who only take evening classes or who only take online classes are most likely not represented by this sample.
- Students should understand that by randomly selecting a sample from the entire population everyone/everything in the population has an equal chance of being selected. This lack of bias allows for a variety of people/objects to be selected so that various characteristics are present in the sample. Therefore, the makeup of the sample is similar

to that of the population, and we can say the sample is representative.

- Instruction includes using technology and/or applets to randomly select a sample from a population to see how the sample has similar characteristics when compared to the population.
- Avoid measurement bias by having an appropriate statistical question. When asking a question to collect data that is representative of the population, the question should be clear, concise and free of any language that may bias the response of any participants.

Common Misconceptions or Errors

- Students may have initially misconceptions regarding how a random sample does not lead to a biased sample. They may incorrectly assume that randomly selecting from the population will inadvertently leave out certain groups. Using simulations and applets can help with this misconception.

Instructional Tasks

Instructional Task 1 (MTR.4.1, MTR.7.1)

A manager for an office supply store has a large shipment of 2,000 calculators being delivered today. If more than 2% of the calculators are defective, he will not accept the shipment and will have to send them all back. Due to time constraints he cannot test all 2,000 calculators to make sure enough of them work; he can only test 100. The manager decides that he will take a sample of 100 calculators and test them to estimate the percentage in the shipment that are defective.

Part A. Why would checking the first 100 calculators most likely not produce data that is representative of the population? Explain.

Part B. What would be the best way to get sample that produces data that is representative of the population? Elaborate on how the manager could get this type of sample.

Part C. Suppose the manager takes a sample using the method that you wrote about in Part B, and he finds that 6 out of 100 calculators are defective and do not work. What should the manager do? Can he feel confident even though he has only sampled 100 calculators? Explain your reasoning.

Instructional Items

Instructional Item 1

A teacher wants a representative sample of students at her school to get feedback on this year's homecoming theme. Which sample would most likely produce data that is representative of the population?

a. She should poll the 30 students in her homeroom class.
b. She should survey 100 students waiting for the buses after school.
c. She should randomly sample 50 students from the entire population of the school.
d. She should talk to each student who eats lunch on campus this upcoming Friday.

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

MA.912.DP.5.3

Benchmark

MA.912.DP.5.3 Compare and contrast sampling methods.

Benchmark Clarifications:
Clarification 1: Instruction includes understanding the connection between probability and sampling methods.
Clarification 2: Sampling methods include simple random, stratified, cluster, systematic, judgement, quota and convenience.

Connecting Benchmarks/Horizontal Alignment

Terms from the K-12 Glossary

- Random sampling

Vertical Alignment

Previous Benchmarks

- MA.6.DP.1.1

Next Benchmarks

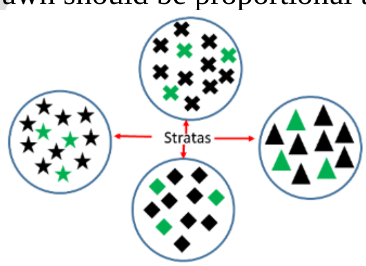
Purpose and Instructional Strategies

In grade 6, students began writing statistical questions in order to collect data for analysis. This idea continued through the grade levels and courses as students collected data to analyze and plot. In Mathematics for College Statistics, students now look at different sampling methods and the pros and cons of such methods.

- Instruction includes the exploration of various sampling methods such as simple random, stratified, cluster, systematic, judgement, quota and convenience. There should also be a discussion of the advantages and disadvantages of each of these sampling methods.
  - **Simple Random Sample:** Simple random sampling is when a sample is selected in such a way that every possible sample of the same size is equally likely to be chosen from the population.

Advantages	Disadvantages
• Highly representative of all subjects participate • Results can be generalized from the sample to the entire population being studied • Can learn about an entire population much faster and more efficiently than collecting data from every member of the population	• Not possible without a complete list of population members • Potentially uneconomical to achieve • Time-scale may be too long; data/sample could change

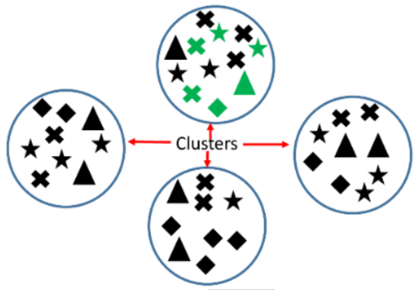
  - **Stratified:** Stratified sampling is a method of sampling that involves the division of a population into smaller sub-groups known as strata. The strata are designed to be homogeneous groups. Then draw a simple random sample from each of these. groups. Samples drawn should be proportional to the size of the strata.



Advantages	Disadvantages
• Can ensure that specific groups are represented, even proportionally, in the	• More complex, requires greater effort than a simple random sample

sample by selecting individuals from strata lists • Less variability	• Strata must be carefully defined

- **Cluster:** Cluster sampling occurs when a random sample is taken from a group or cluster that people belong to. These clusters should be heterogeneous groups. The clusters is chose using a simple random sample and then a simple random sample is taken from that cluster.



Advantages	Disadvantages
• Possible to select randomly when no single list of population members exists, but local lists do • Easy and convenient	• Clusters in a level must be equivalent and some natural ones are not for essential characteristics • Members of units are different from one another, decreasing the techniques effectiveness

- **Systematic:** Systematic sampling is where every member of the population is given a number and then a number is chosen randomly from a list as the first participant. After the first participant, there is an interval that is chosen, for example every 15th person, and every 15th person is chosen on the list.

Advantages	Disadvantages
• Ensures a high degree of representativeness and no need to use a table of random numbers • Allows generalization from the sample to the population being studied • Faster than contacting all members of the population	• Less random than simple random sample • Not possible without a complete list of population members

- **Judgment:** Judgment sampling is where someone uses their own judgment to select participants from the population of interest.

Advantages	Disadvantages
• Inexpensive way of ensuring sufficient numbers of a study	• Can be highly unrepresentative • Vulnerable to errors in judgments that can lead to bias

- **Quota:** Quota sampling is where someone takes a very tailored sample that is in proportion to some characteristic or trait of a population.

Advantages	Disadvantages
• Easy to administer • Fast to create and complete • Inexpensive	• Selection is not random • Selection bias poses a problem • Degree of generalizability is

• Takes into account population proportions • Can be used if probability sampling techniques are not possible	questionable

- **Convenience:** Convenience sampling is when members from a population are chosen based on what is convenient. For example, a student is interested in finding out whether students purchase school lunch and asks students who are in the same classes as they are.

• Inexpensive way of ensuring sufficient numbers of a study	• Can be highly unrepresentative • Degree of generalizability is questionable

- Instruction includes the discussion of the different biases that may occur with these sampling techniques.

Common Misconceptions or Errors

- Students may confuse stratified sampling and cluster sampling.
- Students may have difficulty in understanding how a random sample can be representative of the population.

Instructional Tasks

Instructional Task 1 (MTR.1.1)

Provide students with a sheet of one hundred random rectangles of various areas. The areas of these rectangles should be represented using unit squares. Students will be tasked with estimating the average area of a sample of 5 rectangles using various methods.

Part A. Provide students with these rectangles on a sheet of paper faced down. Tell students to flip the paper over and study the page for 10 seconds. After 10 seconds, they will then flip the paper back over. Students are then to write down their guess as to what the average area of the rectangles are on the sheet.

Part B. Now have students select five rectangles that, in their judgment, are representative of the rectangles that appear on the sheet. Have them write these down and find the average area of the five rectangles that they have chosen.

Part C. Now students will use a random number generator to randomly select five rectangles. Have the students write these five rectangles down and find the average area of the five rectangles that they have chosen.

Part D. On the board, write down the students averages from Part A, Part B and Part C and then create three graphical representations based on students average areas (Guess, Judgment, Random Sample)

- Describe and compare the three distributions.
- What is the mean for each of these distributions?
- How does these compare to the actual average area?

Instructional Items

Instructional Item 1

A high school is discussing the theme for Homecoming. They want to know what theme the students would be interested in. A student decides to group the students as Freshman, Sophomore, Juniors and Seniors and then take a simple random sample from each of these groups to find out which theme the students would like for Homecoming. What type of sampling technique was used?

The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.

MA.912.DP.5.4

Benchmark

MA.912.DP.5.4	Generate multiple samples or simulated samples of the same size to measure the variation in estimates or predictions.

Connecting Benchmarks/Horizontal Alignment

Terms from the K-12 Glossary

- Data
- Population (in data analysis)
- Random sampling

Vertical Alignment

Previous Benchmarks

- MA.6.DP.1.2, MA.6.DP.1.6
- MA.7.DP.1.2
- MA.7.DP.2.4

Next Benchmarks

Purpose and Instructional Strategies

In grade 6, students learned how to calculate the mean and median for numerical data, and they explored how changing values affects the center and variation of numerical data. In grade 7, students experimented with using simulations and comparing the centers and spreads of data. In Mathematics for College Statistics, students combine these middle grades ideas with other benchmarks in this course to examine simulated random samples and naturally occurring sampling variation.

- Instruction relates to benchmark MA.912.DP.5.3.
- Instruction includes students randomly sampling from small populations and large populations in order to compare the differences in statistics. With larger populations, it is recommended that students use technology to generate samples.
- Instruction includes using a variety of random sampling techniques, such as simple random, systematic, cluster and stratified sampling to reinforce previous benchmarks.
- In cluster samples, groups should be heterogeneous. In stratified samples, groups should be homogeneous.
- Students should note the variation in statistics from sample to sample and the variation of sample statistics to the population parameters. As well, learners should consider how to reconcile the differences that are seen in population parameters and sample statistics.
- When generating samples, it is important that the sample size is consistent. This allows for a better comparison and relates to the future statistical topic of standard error.
- Instruction includes a discussion on sampling with and without replacement.

Common Misconceptions or Errors

- Students may incorrectly assume that all samples of the same size from the same population will produce the same statistics.

- Students may incorrectly assume that sample statistics will always match population parameters.
- Students may use convenience samples when generating their own data. These convenience samples can lead to biased results.
- Students may confuse cluster sampling and stratified sampling.

Instructional Tasks

Instructional Task 1 (MTR.7.1)

A teacher is interested in the morning commute times of the students in his homeroom class. The teacher finds out that his students drive, walk, ride or bus from home to school each morning. The results of asking all of his homeroom students "How long did it take you to commute from home to school this morning?" is below. The times are all in minutes.

Student 1 (ride)	12	Student 11 (ride)	15	Student 21 (ride)	15
Student 2 (bus)	20	Student 12 (ride)	20	Student 22 (ride)	10
Student 3 (bus)	15	Student 13 (ride)	5	Student 23 (bus)	20
Student 4 (ride)	5	Student 14 (bus)	30	Student 24 (bus)	18
Student 5 (walk)	3	Student 15 (ride)	35	Student 25 (bus)	20
Student 6 (bus)	25	Student 16 (walk)	15	Student 26 (bus)	5
Student 7 (bus)	10	Student 17 (walk)	3	Student 27 (ride)	20
Student 8 (bus)	30	Student 18 (bus)	30	Student 28 (bus)	20
Student 9 (drive)	25	Student 19 (bus)	15	Student 29 (drive)	30
Student 10 (walk)	10	Student 20 (drive)	10	Student 30 (ride)	5

Consider this homeroom of students as a population. Some parameters are that the mean morning commute time is 16.5 minutes, 10% of students drive, 13.3% of students walk, 43.4% of students bus to school and 33.3% of students ride to school each day.

Part A. Use technology to generate a simple random sample of six students from the population above. Record the sample mean and the proportions of students who drive, walk, bus, and ride to school.

Part B. Use technology to generate another simple random sample six of students from the population above. Record the sample mean and the proportions of students who drive, walk, bus, and ride to school.

Part C. How do the sample means and proportions from the two samples compare?

Part D. Use technology to generate a systematic random sample of six students from the population above. Record the sample mean and the proportions of students who drive, walk, bus and ride to school.

Part E. How do your systematic sample results compare to the results of your first two samples?

Part F. How do all of your sample statistics compare to the population parameters? Is there any variation?

Instructional Items

Instructional Item 1

All seniors planning to attend college in a graduating class at a smaller Florida high school are surveyed to find out their intended to majors. The results are in the table below.

Student 1: Business	Student 15: Business	Student 29: Education	Student 43: Communications	Student 57: Communications	Student 71: Business
Student 2: Education	Student 16: Public Safety	Student 30: Business	Student 44: Education	Student 58: Business	Student 72: Healthcare

Student 3: Business	Student 17: Communications	Student 31: STEM	Student 45: Business	Student 59: Business	Student 73: Public Safety
Student 4: Public Safety	Student 18: STEM	Student 32: STEM	Student 46: Education	Student 60: Education	Student 74: Education
Student 5: Communications	Student 19: STEM	Student 33: Business	Student 47: Healthcare	Student 61: Business	Student 75: Business
Student 6: Business	Student 20: Social Science	Student 34: Business	Student 48: Healthcare	Student 62: Public Safety	Student 76: Business
Student 7: Business	Student 21: Communications	Student 35: STEM	Student 49: Communications	Student 63: Communications	Student 77: Business
Student 8: Healthcare	Student 22: Communications	Student 36: Social Science	Student 50: Communications	Student 64: Business	Student 78: STEM
Student 9: Business	Student 23: Business	Student 37: Business	Student 51: Public Safety	Student 65: Communications	Student 79: Healthcare
Student 10: Healthcare	Student 24: Business	Student 38: Business	Student 52: Social Science	Student 66: Healthcare	Student 80: Public Safety
Student 11: STEM	Student 25: Business	Student 39: Healthcare	Student 53: Social Science	Student 67: STEM	Student 81: Business
Student 12: Business	Student 26: Business	Student 40: Business	Student 54: Business	Student 68: Public Safety	Student 82: STEM
Student 13: Education	Student 27: Business	Student 41: Healthcare	Student 55: STEM	Student 69: Social Science	Student 83: STEM
Student 14: Social Science	Student 28: Business	Student 42: Business	Student 56: Social Science	Student 70: STEM	Student 84: Business

Part A. Use technology to generate a simple random sample of 12 students from the population above. Record the proportion of business majors.

Part B. Again, use technology to generate a simple random sample of 12 students from the population above and record the proportion of business majors.

Part C. How did the proportions of business majors compare from the two samples above? Is there any variation?

The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.

MA.912.DP.5.5

Benchmark

MA.912.DP.5.5 Determine if a specific model is consistent within a given process by analyzing the data distribution from a data-generating process.

Connecting Benchmarks/Horizontal Alignment

- MA.912.DP.1.1, MA.912.DP.1.2
- MA.912.DP.2.1, MA.912.DP.2.4, MA.912.DP.2.5, MA.912.DP.2.7, MA.912.DP.2.9
- MA.912.DP.3.1

Terms from the K-12 Glossary

- Association
- Bivariate data
- Categorical data
- Mean
- Mode
- Random sampling

- Simulation
- Statistical question

Vertical Alignment

Previous Benchmarks

- MA.4.DP.1.3
- MA.5.DP.1.2
- MA.6.DP.1.1, MA.6.DP.1.2
- MA.7.DP.1.3, MA.7.DP.1.5
- MA.8.DP.1.2, MA.8.DP.1.3

Next Benchmarks

Purpose and Instructional Strategies

In previous grades, students have worked with numerical, categorical, univariate, and bivariate data as an introduction to statistical analysis. In Mathematics for College Statistics, students apply this previous knowledge of data analysis to surveys, observational studies and experiments. It is essential to check that the statistical study results in data that is consistent with the model being used. This examination can and should occur in a variety of areas.

- Instruction includes determining if the data generated from a random sample or simulation is displayed in an appropriate manner.
    - For example, bivariate numerical data can be displayed in a scatter plot, while bivariate categorical data can be displayed in a two-way table (or contingency table).
- When instruction centers around Pearson's correlation coefficient and the linear regression equation for numerical bivariate data, students should consistently check that the scatter plot that corresponds with the data is linear. Data displayed in a scatter plot that shows a non-linear pattern should use non-linear regression methods.
- Instruction centering around surveys, observational studies and experiments should include a process of ensuring the correct statistical study is chosen for the outcome that is desired.
- If simulations or experiments are conducted during lessons, it is important to analyze the results to see if they are consistent with what is expected based on prior knowledge about the scenario. Unusual results can lead to conversations about natural random variation and statistical significance.
    - For example, suppose that a teacher says that he has ESP, or the ability read minds and/or the ability to know the future. In order to test his abilities, his students have him guess the colors of 10 randomly drawn cards from a standard deck of 52 cards. If the teacher did not have ESP, we would expect him to guess approximately 50% of the cards' colors correctly (there are two colors – red and black). If he did have ESP, he would be able to get a significantly higher percentage. Suppose he predicts the colors of the cards, and he gets 9 out of 10 colors correct. What should be concluded from the data generated? Does the teacher have ESP? Is the data consistent with the model?
- Instruction includes a discussion on the uncertainty of using samples to make predictions about a population, which is statistical inference. When using sample statistics, we are never 100% certain that we can know the exact value of a population parameter, especially when a population is large. Reconciling natural sampling variation and uncertainty can be difficult, and there are even times when the assertions that are made from samples are incorrect. This can lead into areas such as type I and type II errors in a

college statistics course.

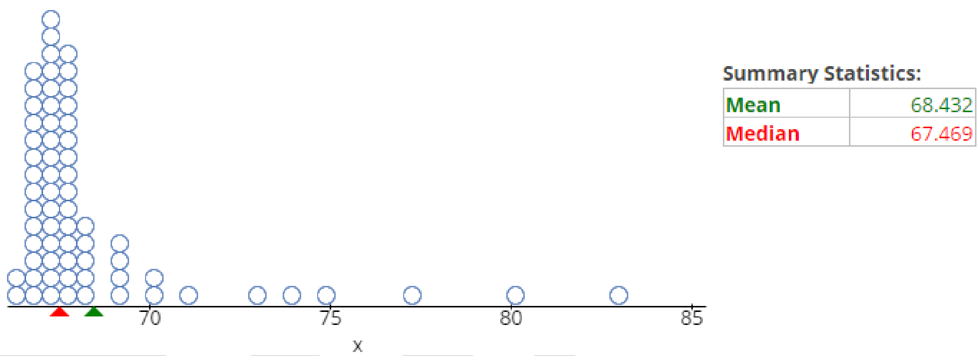
Common Misconceptions or Errors

- Students may be unaccustomed to the overall “messiness” of statistics. There are times where results are inconclusive, or there is not one right answer (like $x = 2$ in algebra) and range of possibilities is discussed.
- Students may incorrectly conclude that the value of a sample statistic has to be equal to the value of a population parameter.
- Students may be in the habit of making calculations without considering whether or not the calculation is appropriate. For example, a student may be asked to find the center of a univariate numerical data set and find the mean; however, when the data set is skewed a median may be more appropriate.

Instructional Tasks

Instructional Task 1 (MTR.4.1, MTR.7.1)

A weather website makes the claim that average daily high temperature in December in Pensacola Florida is 65 degrees Fahrenheit. In order to investigate this claim, a simple random sample of 66 daily high temperatures over the last 10 years is taken. The results are in the dot below.



Part A. What does each dot in the display represent?

Part B. The website claims that the mean daily high temperature in Pensacola is 65 degrees Fahrenheit, but the sample statistic is a mean of around 68.4 degrees Fahrenheit. Is this difference due to natural sampling variation, or does the sample show that the average temperature may have been higher over the last 10 years? Explain your reasoning.

Part C. If another simple random sample of 66 days was taken, do you think we would get the same results as the sample above? Explain.

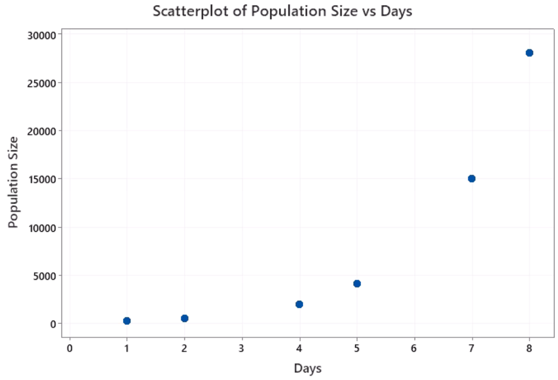
Part D. Would you come to a different conclusion in Part B if the mean for a second sample of 66 days in December in Pensacola had a mean of 67.3 degrees Fahrenheit? Discuss your reasoning with a classmate.

Instructional Items

Instructional Item 1

The size of a population of a particular bacterium is studied and estimated in a petri dish over several days. The corresponding data and scatter plot are below.

Days	1	2	4	5	7	8
Population Size	250	525	2,010	4,100	15,000	28,000



The linear regression equation is calculated to be $Predicted\ Population\ Size = -7809 + 3583(Days)$. Do you feel this linear model accurately describes the population growth of the bacteria? Explain your reasoning.

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

MA.912.DP.5.6

Benchmark

MA.912.DP.5.6	Determine the appropriate design, survey, experiment or observational study, based on the purpose. Articulate the types of questions appropriate for each type of design.

Connecting Benchmarks/Horizontal Alignment

- MA.912.DP.1.1, MA.912.DP.1.2

Terms from the K-12 Glossary

- Association
- Data
- Population (in data analysis)
- Random sampling
- Repeated experiment
- Statistical question

Vertical Alignment

Previous Benchmarks

- MA.6.DP.1.1
- MA.7.DP.1.1, MA.7.DP.1.5
- MA.7.DP.2.1
- MA.8.DP.2.2, MA.8.DP.2.3

Next Benchmarks

Purpose and Instructional Strategies

In middle grades, students examined the processes for generating data and making calculations from data, and they continued these processes to a degree in Algebra I. In Mathematics for College Statistics, students learn about appropriate statistical design for the first time. While some terminology may be unfamiliar to students, these ideas can be related to familiar topics.

- When determining the design of a statistical study, students should first decide the type of conclusion that will be made. In order to infer a causal relationship, the design should be based on an experimental study. In order to show a link or an association, the design can be based on an observational study.

- Instruction includes that observational studies should include random sampling in order to select participants, and participants in an observational study are typically given a survey or are asked questions about their opinions or events that have already occurred. Again, an observational study can lead to a conclusion that there is a link or an association between variables.
- Instruction includes experimental studies that are designed such that participants are randomly assigned to different groups (at least two or more). Different aspects to be discussed with experimental studies include treatment groups, control groups, placebos, blinding and double-blinding. An experimental study can lead to a conclusion that there is a cause-and-effect relationship among variables.
    - A treatment group is a group in an experiment that receives the treatment of interest.
    - A control group is a group in an experiment that does not receive the treatment of interest.
    - A placebo is something that may be similar to the actual treatment but should have no effect on the subjects taking the placebo.
    - Blinding occurs when the subjects in the experiment are not aware of which group they are in, or even that other groups exist in the experiment.
    - Double-blinding occurs when the subjects in the experiment are not aware of which group they are in, nor do the individual collecting data know which group the subject is in.
- Instruction includes discussions on what questions should be asked, how questions should be asked, the data that will be generated, how this data will be generated and ethics that may be involved in certain observational studies or experimental studies. There should also be a discussion of the advantages and disadvantages of each of these types of statistical study designs.
- Avoiding measurement bias, sampling bias and the presence of nonresponse bias in surveys should be addressed during instruction.
- Instruction includes identifying the benefits of large sample sizes.

Common Misconceptions or Errors

- Students may have difficulty when determining which type of study may be more appropriate for a particular statistical question.
- Students may have difficulty when identifying and including all of the necessary aspects in designing a study.

Instructional Tasks

Instructional Task 1 (MTR.2.1, MTR.7.1)

Suppose that a college professor would like to show that there is a link between using the on-campus tutoring center and earning passing grades in college algebra classes taught at her college.

Part A. What type of statistical study should the professor use? Explain.

Part B. How should she design her study? Be specific in the elements that should be present.

Part C. What questions should be asked of participants, and what type(s) of data would be produced?

Part D. Suppose that the data show that students who utilize the on-campus tutoring center three or more times in a semester are more likely to earn a passing grade in college algebra. What conclusion can be made?

Part E. What if the professor wants to show that utilizing the on-campus tutoring center leads to earning passing grades in college algebra; would this study be able to show this result? What changes, if any, would need to be made, and what elements would need to be added in to the statistical study?

Instructional Items

Instructional Item 1

Suppose a pharmaceutical company wants to show that a new medication will reduce the number of migraines in patients who suffer from severe migraines. How could the company design a statistical study to possibly show this? What questions should be asked, and what conclusions could be drawn?

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

MA.912.DP.5.7

Benchmark

MA.912.DP.5.7 Compare and contrast surveys, experiments and observational studies.

Benchmark Clarifications:
Clarification 1: Instruction includes understanding how randomization relates to sample surveys, experiments and observational studies.

Connecting Benchmarks/Horizontal Alignment

Terms from the K-12 Glossary

- Data
- Population (in data analysis)
- Random sampling
- Association
- Repeated experiment
- Statistical question

Vertical Alignment

Previous Benchmarks

- MA.6.DP.1.1
- MA.7.DP.1.1, MA.7.DP.1.5
- MA.7.DP.2.1
- MA.8.DP.2.2, MA.8.DP.2.3
- MA.912.DP.1.4

Next Benchmarks

Purpose and Instructional Strategies

In middle grades, students examined the processes for generating data and performing calculations from data. Students continued these processes to a degree in Algebra I. In Mathematics for College Statistics, students learn to compare and contrast surveys, observations, and experiments. This benchmark has a direct relationship to the benchmark MA.912.DP.5.6.

- Instruction defines a survey to be the set of questions asked to participants. Surveys can be deployed in an observation or an experiment.

Advantages	Disadvantages
• Can be administered in a variety of ways such as by telephone, mail, online, mall interviews,	• Are dependent upon the respondent’s honesty and motivation when answering (can create a

etc. • Are efficient for collecting data from a large populations. • Can be designed to focus only on the needed response questions. • Are applicable to a wide range of topics.	response bias). • Can be flawed by non-response bias (no response given). • Can possess questions or answer choices that may be interpreted differently by different respondents.

- Students should have practice examining and comparing statistical studies; students should examine how participants were selected or organized. If subjects are randomly sampled and all asked the same set of questions, an observational study was utilized. If the study was designed such that participants are randomly assigned to different groups (at least two or more), an experimental study was utilized.

	• More Validity • Can determine causality • Randomized and blinded	• May require less resources and/or time • Less ethical concerns when dealing with potentially harmful exposures • Good if outcome of interest is rare
	• May require more resources and/or time • Ethical concerns for certain exposures • Difficult if outcome being studied is rare	• Less validity → difficult to determine causality • No randomization or blinding

- For enrichment of this benchmark, different aspects can be discussed with experimental studies, including treatment groups, control groups, placebos, blinding and double-blinding.
    - A single-blinded experiment is designed so that the participants are not aware of the treatment being administered in the study. A double-blinded experiment is designed so that participants and those collecting data are not aware of the treatment being administered in the study.
- As an expectation of this benchmark, students examine different types of studies and decide the conclusion that can be made.
    - For example, if the study is experimental, it is possible to infer a cause and effect relationship between variables.
    - If the study is observational, do not conclude a causal relationship. Instead, infer that there is a link or an association between variable.
- Instruction includes emphasizing that observational studies should include random sampling, while experimental studies should include random assignment among the various groups.

Common Misconceptions or Errors

- Students may initially make incorrect assumptions when determining if a study is an observation or experiment.
- Students may have difficulty deciding upon an appropriate conclusion from the statistical study.
- Students may believe experiments must use a placebo or have a control group.

Instructional Tasks

Instructional Task 1 (MTR.4.1, MTR.7.1)

Suppose a nutritionist is interested in the relationship between coffee consumption and a person's sleeping habits.

Part A. Explain, in detail, how to design a statistical study that could possibly show a link between increased coffee consumption and a decrease in an individual's average number of hours slept per night. Be sure to include the use of a survey and use proper statistical terminology.

Part B. Explain, in detail, how to design a statistical study that could possibly show increased coffee consumption causes a decrease in an individual's average number of hours slept per night. Be sure to include the use of a survey and use proper statistical terminology.

Part C. Partner with a classmate to compare and contrast the designs of the two studies above.

Instructional Items

Instructional Item 1

Mr. Artemis is reviewing his students' grades and their attendance over the course of the school year. He notices that students who miss very few classes have higher grades, and that as students miss more and more classes, their grades tend to decrease.

Part A. Should Mr. Artemis conclude that missing fewer classes will cause students' grades to increase? Explain why this is or is not an appropriate conclusion.

Part B. Would it be possible (or ethical) to design an experimental study where Mr. Artemis could determine if missing fewer classes will cause students' grades to increase? Explain.

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

MA.912.DP.5.11

Benchmark

MA.912.DP.5.11	Evaluate reports based on data from diverse media, print and digital resources by interpreting graphs and tables; evaluating data-based arguments; determining whether a valid sampling method was used; or interpreting provided statistics.

Example: A local news station changes the y-axis on a data display from 0 to 10,000 to include data only within the range 7,000 to 10,000. Depending on the purpose, this could emphasize differences in data values in a misleading way.

Benchmark Clarifications:

Clarification 1: Instruction includes determining whether or not data displays could be misleading.

Connecting Benchmarks/Horizontal Alignment

- MA.912.DP.1.1, MA.912.DP.1.2

Terms from the K-12 Glossary

- Axes
- Bar graph
- Bivariate data
- Box plot
- Categorical data
- Circle graph
- Data
- Histogram

- Line graph
- Line plot
- Random sampling
- Repeated experiment
- Scatter plot
- Stem-and-leaf plot

Vertical Alignment

Previous Benchmarks

- MA.6.DP.1.5
- MA.7.DP.1.4, MA.7.DP.1.5
- MA.8.DP.1.1

Next Benchmarks

Purpose and Instructional Strategies

In earlier grade levels, students created and analyzed data for various situations. These were either student gathered data or data from outside resources. In Mathematics for College Statistics, students continue this exploration by analyzing different media resources. In addition to this, students begin to see how statistics can be used to mislead.

- Instruction includes the use of various media resources such as newspapers, magazines, internet, etc. Students should be able to identify whether a valid sampling method was used to collect data.
- Instruction includes the exploration of misleading graphs, and why they are misleading. Some types of misleading graphs are those that use a vertical axis that does not start at 0, use inconsistent scales or use different shapes or images instead of bars in bar charts. Students should also discuss what the impact misleading graphs and data may have on society when it is allowed to be published as fact.
- When looking at graphs that are misleading, students should have experience in deciding what can be done, if anything, to correct the presentation of the graph so that it is more accurate.

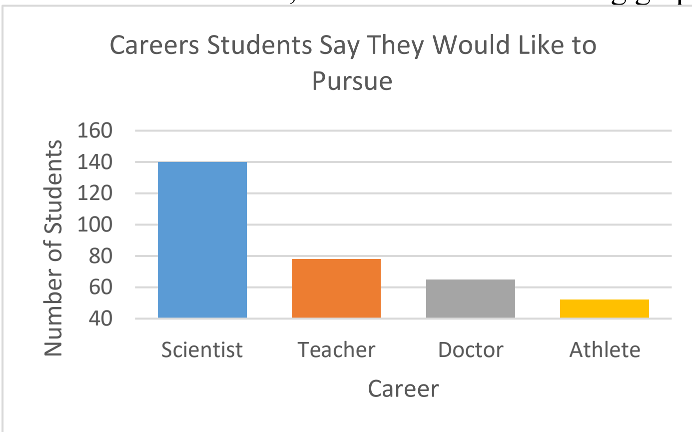
Common Misconceptions or Errors

- Students may automatically look at the image of the graph without taking into consideration anything else. Be sure that students look beyond the surface and identify the following:
  - Is there a scale to the graph and where does it begin? If there is not a scale, can we really trust the data?
  - Is the scale consistent throughout?
  - If a circle graph (pie chart) is used, do the percentages add up to be 100%?
  - If it is a pictograph, are the images out of proportion for what they represent?
- Students may assume that the results of a given graph, table or conclusion that is given in media is automatically true. The purpose of this class it to empower students to decipher statistics in the world around them and not accept things at face value. This will gain them the knowledge that they need in order to become informed consumers.

Instructional Tasks

Instructional Task 1 (MTR.6.1, MTR.7.1)

Jaylena did a survey asking students, “What career do you want to pursue after you graduate high school?” Once she collected her data, she created the following graph.



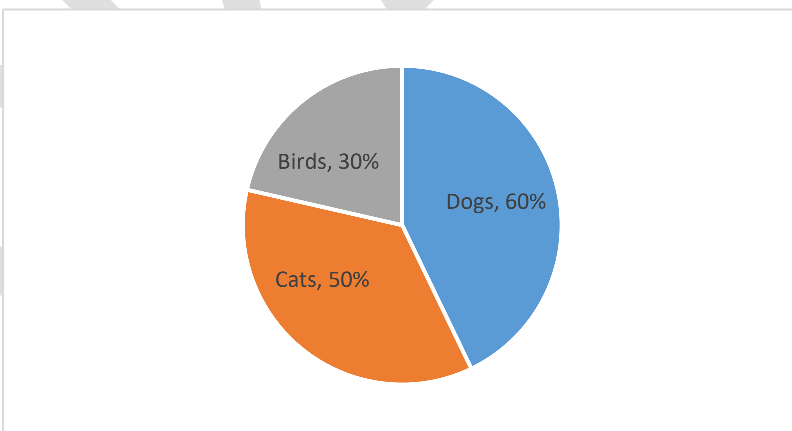
She states that, “There are a huge number of students who are interested in pursuing a career as a scientist – way more than any other career!” Jeremy sees Jaylena’s graph and says, “Your numbers may be right, but your graph is misleading.”

Do you agree with Jeremy that the graph is misleading? If you agree, how is it misleading and in what ways can it be fixed. If you disagree, why is it not misleading?

Instructional Items

Instructional Item 1

The following is a circle graph showing the percentage of students who own dogs, cats and birds.



Is the given graph a valid representation of the data? Explain why or why not.

*The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.

MA.912.LT.5 *Apply properties from Set Theory to solve problems.*

MA.912.LT.5.4

Benchmark

MA.912.LT.5.4 Perform the set operations of taking the complement of a set and the union, intersection, difference and product of two sets.

Benchmark Clarifications:
Clarification 1: Instruction includes the connection to probability and the words AND, OR and NOT.

Connecting Benchmarks/Horizontal Alignment

- MA.912.DP.4.1, MA.912.DP.4.2, MA.912.DP.4.3, MA.912.DP.4.4

Terms from the K-12 Glossary

- Event
- Sample space

Vertical Alignment

Previous Benchmarks

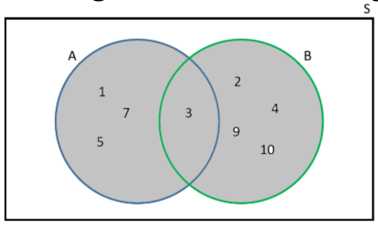
- MA.7.DP.2.1, MA.7.DP.2.4
- MA.8.DP.2.1, MA.8.DP.2.2, MA.8.DP.2.3

Next Benchmarks

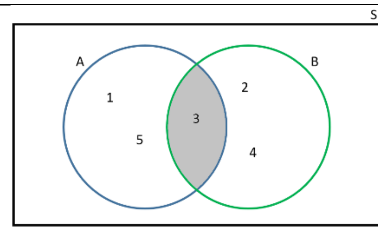
Purpose and Instructional Strategies

In grades 7 and 8, students began exploring probability for both simple and compound events. They identified the sample space for these in order to explore theoretical and experimental probabilities. In Mathematics for College Statistics, students expand on this idea by going into set theory and how this can be used when exploring probabilities.

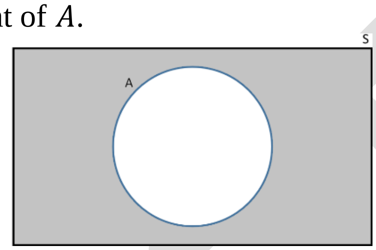
- The instruction of this benchmark and MA.912.LT.5.5 should be taught together as the Venn Diagrams are beneficial for students to have an understanding of set notation and how to interpret set notation operations.
- Instruction includes the use of diagrams in order to explore and illustrate concepts of complements, unions, intersections, difference and products of two sets.
    - The union of two sets is a set containing all elements that are in A or in B.
        - For example, if $A = \{1, 3, 5, 7\}$ and $B = \{2, 3, 4, 9, 10\}$, then $A \cup B$ is $\{1, 3, 5, 7\} \cup \{2, 3, 4, 9, 10\} = \{1, 2, 3, 4, 5, 7, 9, 10\}$. The Venn Diagram of this is the following where the shaded region represents $A \cup B$:



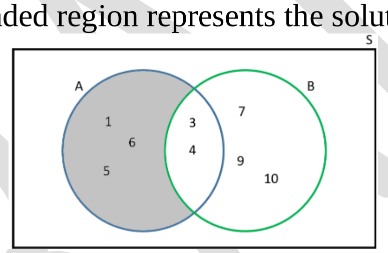
    - The intersection of two sets is denoted as $A \cap B$ and consists of all elements that are both in A and B.
        - For example, if $A = \{1, 3, 5\}$ and $B = \{2, 3, 4\}$, then $A \cap B$ is $\{1, 3, 5\} \cap \{2, 3, 4\} = \{3\}$. The Venn Diagram of this is the following where the shaded region represents $A \cap B$:



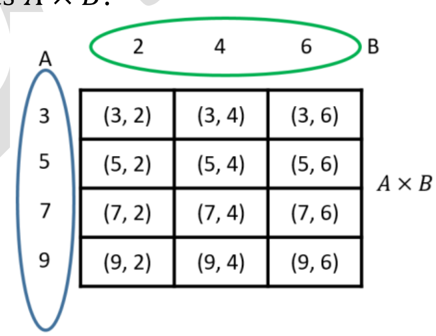
- The complement of a set can be denoted as A^c, $\bar{A}$, A' or $\sim A$. This is the set of all elements that are in the universal set S but not in A.
  - The Venn Diagram below shows that the shaded region represents the complement of A.



- The difference, $A - B$, consists of elements in A but not in B. This can also be seen as $A - B = A \cap B^c$.
  - For example, if $A = \{1, 3, 4, 5, 6\}$ and $B = \{3, 4, 7, 9, 10\}$, then $A - B$ is $\{1, 3, 4, 5, 6\} - \{3, 4, 7, 9, 10\} = \{1, 5, 6\}$. The Venn Diagram below show that the shaded region represents the solution to $A - B$.

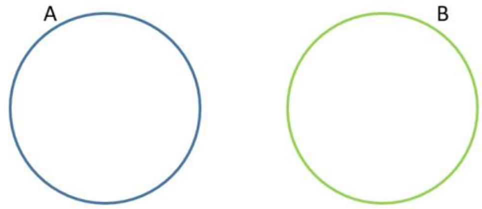


- The product of sets A and B is the set of ordered pairs. This can be represented as $A \times B$. This is also known as Cartesian Product.
  - For example, $A = \{3, 5, 7, 9\}$ and $B = \{2, 4, 6\}$, so $A \times B$ can be written as $\{3, 5, 7, 9\} \times \{2, 4, 6\}$, which equals $\{(3, 2), (3, 4), (3, 6), (5, 2), (5, 4), (5, 6), (7, 2), (7, 4), (7, 6), (9, 2), (9, 4)\}$. The following diagram illustrates $A \times B$.



- Instruction includes connecting set notation with AND, OR and NOT.
  - The event $A \cup B$ is the event that "A or B occurs."
  - The event $A \cap B$ is the event that "A and B occurs."
  - The event A^c (complement of A) is the event that "A does not occur."

- In MA.912.DP.4.7, students explored mutually exclusive events. In this benchmark, instruction includes the use of mutually exclusive (also called disjoint) events.
    - The Set A and Set B are mutually exclusive (disjoint). The following is a Venn Diagram that represents this:



Common Misconceptions or Errors

- Students may confuse the notation that is used for union ($\cup$) and intersection ($\cap$).
- Students may confuse that union is related to OR and intersection is related to AND.
- When finding the product of sets, students may actually multiply these values instead of writing them as coordinate pairs.
- When finding the difference of sets, students may actually try to subtract these values instead of seeing as A intersecting with the complement of B. Using a visual model may assist students with this concept.
- When given a Venn Diagram and are asked to find the union, students may omit the intersection in their solution.

Instructional Tasks

Instructional Task 1 (MTR.7.1)

The numbers 1 through 25 are written on slips of paper that are then placed in a box. Students draw a slip to determine the order in which they will give a presentation. Event A is being one of the first 10 students to give their presentation. Event B is picking a multiple of 4.

Part A. What is the same space for Event A and Event B?
Part B. Find $P(A)$. What does this represent?
Part C. What notation would be used to find the probability of A or B? Find this probability.
Part D. What does $P(A \cap B)$ mean? Find this probability.
Part E. What is the complement of A?

Instructional Items

Instructional Item 1

You roll a ten-sided die. Event A is rolling an even number and Event B is rolling a number that is prime.

Part A. What is the sample space for Event A and Event B?
Part B. Find the following: $A \cup B$; $A \cap B$; A^c

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*

MA.912.LT.5.5

Benchmark

MA.912.LT.5.5 Explore relationships and patterns and make arguments about relationships between sets by using Venn Diagrams.

Connecting Benchmarks/Horizontal Alignment

- MA.912.DP.4.1, MA.912.DP.4.2, MA.912.DP.4.3, MA.912.DP.4.4

Terms from the K-12 Glossary

- Event
- Sample space

Vertical Alignment

Previous Benchmarks

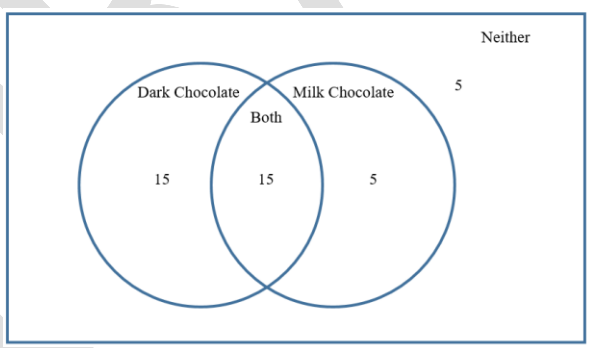
- MA.7.DP.2.1, MA.7.DP.2.4
- MA.8.DP.2.1, MA.8.DP.2.2, MA.8.DP.2.3

Next Benchmarks

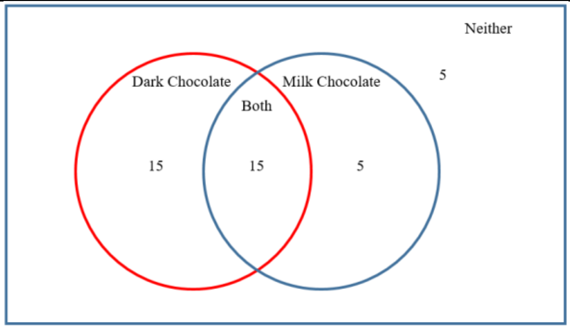
Purpose and Instructional Strategies

In grades 7 and 8, students began exploring probability for both simple and compound events. They identified the sample space for these in order to explore theoretical and experimental probabilities. In Mathematics for College Statistics, students continue this exploration through Venn Diagrams.

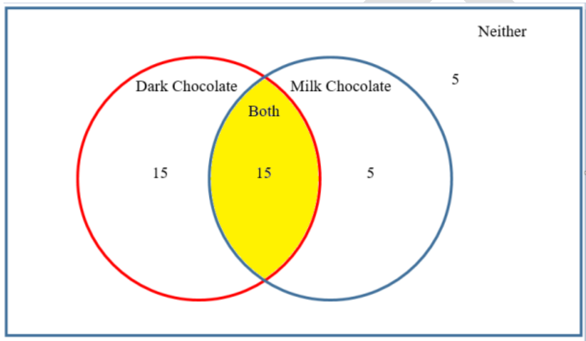
- Instruction of this benchmark should be done in conjunction with MA.912.LT.5.4.
- Students should be able to translate a context into a Venn diagram and then use the diagram to answer questions.
- Instruction includes the use of a sample space to create a Venn diagram. Then move towards using counts and probabilities.
- For example, a survey about the whether someone likes dark chocolate or milk chocolate was given to 40 students. The survey showed that 30 liked dark chocolate, 20 liked milk chocolate, 15 liked both and 5 liked neither. This is the Venn Diagram that represents this situation:



- Instruction includes the use of probability notation in which the students would then need to interpret it using a Venn Diagram.
  - Using the previous example about chocolate, what is the $P(Dark\ Chocolate \cup Milk\ Chocolate)$? The $P(Dark\ Chocolate \cup Milk\ Chocolate) = P(Dark\ Chocate) + P(Milk\ Chocolate) - P(Dark\ Chocolate\ and\ Milk\ Chocolate)$. So by looking at the Venn Diagram, the area in the red circle represents Dark Chocolate. So, $P(Dark\ Chocolate) = \frac{30}{40} = \frac{3}{4}$.



The Milk Chocolate is represented in the blue circle. So, the $P(Milk\ Chocolate) = \frac{20}{40} = \frac{1}{2}$. Now, $P(Dark\ Chocolate \cap Milk\ Chocolate)$ can be seen below highlighted in yellow. This probability is $\frac{15}{40}$ or $\frac{3}{8}$.



Now it can be solved as $P(Dark\ Chocolate \cup Milk\ Chocolate) = \frac{3}{4} + \frac{1}{2} - \frac{3}{8} = \frac{9}{8}$.

Common Misconceptions or Errors

- When working with Venn Diagrams, students may forget to account for the intersection when creating their diagrams.
- Students may not consider elements that are not within the circles of the Venn Diagram when interpreting data.

Instructional Tasks

Instructional Task 1 (MTR.2.1, MTR.6.1, MTR.7.1)

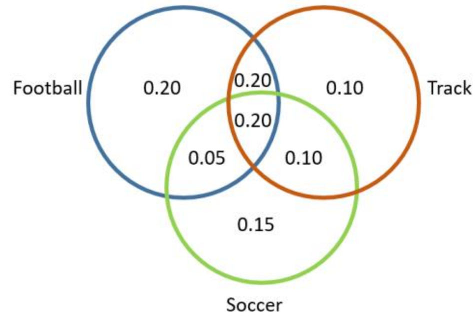
100 students were surveyed as to whether they owned a dog or a cat. Of the students, 45 stated they owned a dog, 20 stated they owned a cat and 15 stated that they owned both.

Part A. Create a Venn diagram to represent this situation.
Part B. How many students did not own a dog or a cat? How do you know?
Part C. What is the probability that a student owns a dog or a cat? Explain how you found this probability.
Part D. What is the probability that a student owns a cat but not a dog? Explain how you found this probability.
Part E. Jana states that she found $P(dog \cap cat)$ by doing the following: $45 + 20 + 15 = 80$. So she says that 80 people own a dog or a cat. Do you agree? Explain why or why not.

Instructional Items

Instructional Item 1

Students were surveyed as to whether they participated in football, soccer or track. The following is the Venn diagram from this survey.



Part A. What percentage of students stated that they participated in football?
Part B. What percentage of students stated that they participated in track but not soccer?
Part C. What percentage of students stated that they participated in all three sports?

**The strategies, tasks and items included in the B1G-M are examples and should not be considered comprehensive.*