VIRGINIA BOARD OF EDUCATION
AGENDA ITEM

Agenda Item: **G**

Date: **March 17, 2022**

Title: **Final Review of the Proposed *Data Science Standards of Learning* and *Data Science Standards of Learning Curriculum***

Presenter: **Tina Mazzacane, K-12 Mathematics Coordinator**

Email: **Tina.Mazzacane@doe.virginia.gov** **Phone: (804) 225-4849**

Purpose of Presentation:
Action required by state or federal law or regulation.

Executive Summary:
The proposed *Data Science Standards of Learning (SOL)* and proposed *Data Science Standards of Learning Curriculum Framework* are being presented for a final review to the Virginia Board of Education (Board). The proposed *Data Science Standards of Learning and Curriculum Framework* may be found in **Attachments A and B**, respectively. On November 18, 2021, the Board received for the first review the proposed *Data Science SOL* and *Curriculum Framework.*

The proposed *Data Science Standards of Learning* and proposed *Data Science Standards of Learning Curriculum Framework* would be an addition to the current 2016 *Mathematics Standards of Learning*. Standards for data science do not currently exist.

The following summarizes the actions taken since November 18, 2021 when the Board received the proposed *Data Science SOL* and *Curriculum Framework* for first review.

- The Virginia Department of Education (VDOE) received public comment from various stakeholders regarding the Proposed Revised *Data Science SOL* and *Curriculum Framework* from December 17, 2021-January 31, 2022. The stakeholders providing public feedback included parents, educators, representatives from business and industry, and other community members. Feedback was accepted at two public hearings and through email and electronic submission to the Virginia Department of Education. The public hearings were presided over by Board members and held on January 11, 2022 and January 20, 2022.

    - Public comment was received from three individuals at the public hearing held January 11, 2022.
    - Public comment was received from ten individuals at the public hearing held January 20, 2022.
    - Written feedback was received from 16 stakeholders including parents, educators, business and industry representatives, and community members.

- The VDOE convened an external virtual review meeting on February 7, 2022, with educators from state institutes of higher education, state mathematics organizations, and business and industry representatives to review and seek feedback regarding the proposed *Data Science Standards of Learning*. Invitations were extended to:
    - 25 educators from state institutes of higher education;
    - 7 representatives from state mathematics organizations; and
    - 35 business and industry representatives.

- VDOE staff met with the proposed *Data Science Standards of Learning* lead developers during February 2022 to review the comments and feedback collected and make edits to the proposed *Data Science Standards of Learning*.

The attached drafts of the proposed *Data Science Standards of Learning* and the proposed *Data Science Standards of Learning Curriculum Framework* include revisions since November 18, 2021, when the Board received the proposed *Data Science SOL* and *Curriculum Framework* for first review in response to public comment, as listed.

- Edits to provide clarity and consistency in language;
- Edits to specify and expound upon the application of statistics and mathematics; and
- Edits to focus more on data bias versus ethics.

All edits found in the proposed drafts have been tracked using the following system:

- a single underline (<u>sample</u>) indicates content added to the initially Proposed *Data Science Standards of Learning* or *Curriculum Framework*; and
- a single strikethrough (~~sample~~) indicates content deleted from the initially Proposed *Data Science Standards of Learning* or *Curriculum Framework*.

The proposed *Data Science Standards of Learning* and proposed *Data Science Standards of Learning Curriculum Framework* also align to the Virginia Board of Education Goal 2 (Rigorous Standards to Promote College and Career Readiness) from the 2018-2023 Comprehensive Plan. Goal 2 states, "the Virginia Board of Education has made a commitment to maintain rigorous and relevant expectations for students that meet or exceed national and international benchmarks for college and career readiness."

Action Requested:
Final review: Action requested at this meeting.

Superintendent's Recommendation:

Rationale for Action:
The proposed *Data Science Standards of Learning* and *Curriculum Framework* will support school divisions in offering a rigorous high school course in Data Science which will provide students with an option to earn a half credit (semester) or one credit (year-long) in mathematics toward graduation starting in 2022-2023. *Mathematics Standards of Learning* focused on Data Science do not currently exist and these standards will expand mathematics learning opportunities and support stronger data literacy skills for students.

Previous Review or Action:
Date: November 18, 2021
Action: First Review

Date: January 11, 2022
Action: Public Hearing

Date: January 20, 2022
Action Public Hearing

Background Information and Statutory Authority:
Data science is a growing field that allows for the analysis of data through the application of mathematics, statistics, computer science, and information technology. The demand for data science knowledge and skills permeates more and more careers. To be productive and thoughtful citizens, children must learn to be discerning consumers of data. Preparing students to be data-literate citizens who can navigate a world that is inundated with data requires rethinking mathematics education. A high school course in data science will provide students with an understanding of how to visualize and interpret data, identify potential bias in data, and leverage data as a tool to support change and innovation. The standards support problem solving using large data sets through an inquiry-based approach.

The proposed *Data Science Standards of Learning* and proposed *Data Science Standards of Learning Curriculum Framework* align to the *Profile of a Virginia Graduate,* which describes the knowledge, skills, competencies, and experiences students should attain during their K-12 education to make them "life-ready," and prepared to succeed in the evolving economy. In a course based on the *Data Science SOL,* students will explore content through critical thinking, creative thinking, collaboration, communication, and citizenship. In addition to addressing the "5 C's," the proposed *Data Science Standards of Learning* and proposed *Data Science Standards of Learning Curriculum Framework* support the *Profile of a Virginia Graduate* goal of establishing multiple paths toward college and career readiness for students.

The *Code of Virginia* requires Board of Education to establish educational objectives and review the Standards of Learning periodically, as referenced below:

> *Code of Virginia* § 22.1-253.13:1
>
> B. "The Board of Education shall establish educational objectives known as the Standards of Learning, which shall form the core of Virginia's educational program, and other educational objectives, which together are designed to ensure the development of the skills that are necessary for success in school and for preparation for life in the years beyond. At a minimum, the Board shall establish Standards of Learning for English, mathematics, science, and history and social science. The Standards of Learning shall not be construed to be regulations as defined in § 2.2-4001."
>
> "The Board shall seek to ensure that the Standards of Learning are consistent with a high-quality foundation educational program. The Standards of Learning shall include, but not be limited to, the basic skills of communication (listening, speaking, reading, and writing); computation and critical reasoning, including problem solving and decision making; proficiency in the use of computers and related technology; computer science and computational thinking, including computer coding; and the skills to manage personal finances and to make sound financial decisions."
>
> "The Standards of Learning in all subject areas shall be subject to regular review and revision to maintain rigor and to reflect a balance between content knowledge and the application of knowledge in preparation for eventual employment and lifelong learning. The Board of Education shall establish a regular schedule, in a manner it deems appropriate, for the review and revision as may be necessary, of the Standards of Learning in all subject areas. Such review of each subject area shall occur at least once every seven years. Nothing in this section shall be construed to prohibit the Board from conducting such review and revision on a more frequent basis."

New academic content Standards of Learning for mathematics were first developed in 1995. Pursuant to legislation from the 2000 Virginia General Assembly, the Boardestablished a seven-year cycle for review of the Standards of Learning. As a result, the 1995 *Mathematics Standards of Learning* were reviewed in 2001, 2009, and 2016.

Timetable for Further Review/Action:

Upon final approval by the Board of the proposed *Data Science Mathematics Standards of Learning* and the proposed *Data Science Standards of Learning Curriculum Framework*, the VDOE will post a final version on the website.

Data Science courses, based on the *Data Science Standards of Learning*, may be offered as a semester (½ credit) or year-long (1 credit) in mathematics toward graduation starting in 2022-2023. Upon final approval, Data Science will be brought to the Board as an addition to the Board of Education Approved Courses to Satisfy Graduation Requirements for the Standard, Advanced Studies, and Modified Standard Diplomas in Virginia Public Schools document. Locally-developed courses based on the *Data Science SOL* will be considered at or above the level of Algebra II. There will be no assumed prior knowledge of computer science or coding prior to taking a course based on the *Data Science Standards of Learning*. The VDOE will support school divisions in making local decisions about prerequisite courses and possible course trajectories. Teachers for this course must be certified in Mathematics (3100), Computer Science (2004), or Computer Science Specialist (3010).

Pending final approval of the proposed *Data Science Standards of Learning* and proposed *Data Science Standards of Learning Curriculum Framework* by the Board, a small pilot implementation of Data Science with a limited number of school divisions during the 2022-2023 school year is planned. Teachers involved in the pilot will receive professional development and participate in a professional learning cohort. Teams of teachers and specialists will work to develop draft curricular resources that will support instruction in Data Science. The VDOE is also working with Virtual Virginia in creating a pilot course that can be offered in 2022-2023. Full implementation of the *Data Science Standards of Learning* and *Data Science Standards of Learning Curriculum Framework* will occur during the 2023-2024 school year.

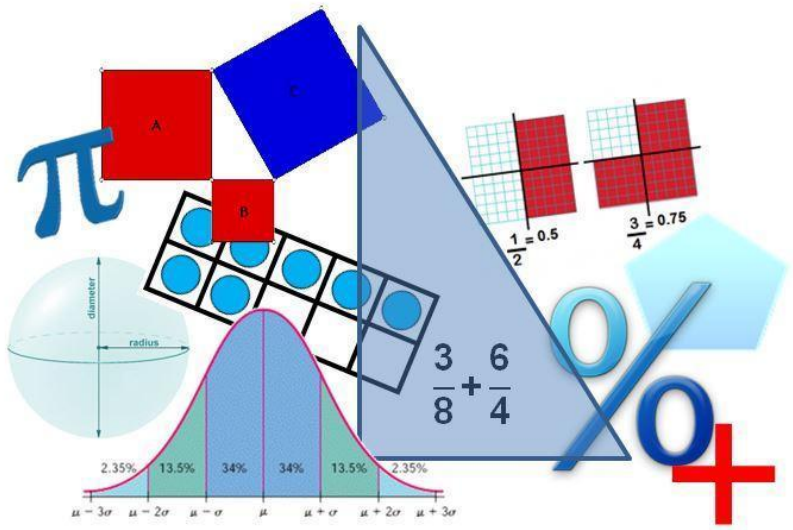
The Implementation Timeline included in **Attachment C** includes details regarding implementation.

Impact on Fiscal and Human Resources:
The implementation of the *Data Science Standards of Learning* and *Data Science Standards of Learning Curriculum Framework* along with professional learning for the pilot implementation and the development of draft resources can be absorbed by the agency's existing resources at this time. If the agency is required to absorb additional responsibilities related to this activity, other services may be impacted. School divisions may be impacted by providing release time for teachers to participate in professional learning.

Mathematics

2016 Standards of Learning

Data Science
Curriculum Framework



Board of Education
Commonwealth of
Virginia

Superintendent of Public Instruction
~~James F. Lane~~Jillian A. Balow

~~**Chief Academic Officer/**~~**Assistant Superintendent for** ~~**Instruction**~~**Learning and Innovation**
Michael Bolling

Office of STEM and Innovation
Brendon Albon, STEM and Innovation Director
Tina Mazzacane, Mathematics Coordinator
Kristin Williams-Faus, Mathematics and Special Education Specialist
Debra Delozier, Mathematics Specialist

NOTICE
The Virginia Department of Education does not unlawfully discriminate on the basis of race, color, sex, national origin, age, or disability in employment or in its educational programs or services.

Virginia 2016 *Mathematics Standards of Learning Curriculum Framework*

Introduction

The 2016 *Mathematics Standards of Learning Curriculum Framework*, a companion document to the 2016 *Mathematics Standards of Learning*, amplifies the *Mathematics Standards of Learning* and further defines the content knowledge, skills, and understandings that are measured by the Standards of Learning assessments. The standards and *Curriculum Framework* are not intended to encompass the entire curriculum for a given grade level or course. School divisions are encouraged to incorporate the standards and *Curriculum Framework* into a broader, locally designed curriculum. The *Curriculum Framework* delineates in greater specificity the minimum content that all teachers should teach and all students should learn. Teachers are encouraged to go beyond the standards as well as to select instructional strategies and assessment methods appropriate for all students.

The *Curriculum Framework* also serves as a guide for Standards of Learning assessment development. Students are expected to continue to connect and apply knowledge and skills from Standards of Learning presented in previous grades as they deepen their mathematical understanding. Assessment items may not and should not be a verbatim reflection of the information presented in the *Curriculum Framework*.

Each topic in the 2016 *Mathematics Standards of Learning Curriculum Framework* is developed around the Standards of Learning. The format of the *Curriculum Framework* facilitates teacher planning by identifying the key concepts, knowledge, and skills that should be the focus of instruction for each standard. The *Curriculum Framework* is divided into two columns: Understanding the Standard and Essential Knowledge and Skills. The purpose of each column is explained below.

Understanding the Standard
This section includes mathematical content and key concepts that assist teachers in planning standards-focused instruction. The statements may provide definitions, explanations, examples, and information regarding connections within and between grade level(s)/course(s).

Essential Knowledge and Skills
This section provides a detailed expansion of the mathematics knowledge and skills that each student should know and be able to demonstrate. This is not meant to be an exhaustive list of student expectations.

Mathematical Process Goals for Students

The content of the mathematics standards is intended to support the following five process goals for students: becoming mathematical problem solvers, communicating mathematically, reasoning mathematically, making mathematical connections, and using mathematical representations to model and interpret practical situations. Practical situations include real-world problems and problems that model real-world situations.

Mathematical Problem Solving

Students will apply mathematical concepts and skills and the relationships among them to solve problem situations of varying complexities. Students also will recognize and create problems from real-world data and situations within and outside mathematics and then apply appropriate strategies to determine acceptable solutions. To accomplish this goal, students will need to develop a repertoire of skills and strategies for solving a variety of problems. A major goal of the mathematics program is to help students apply mathematics concepts and skills to become mathematical problem solvers.

Mathematical Communication

Students will communicate thinking and reasoning using the language of mathematics, including specialized vocabulary and symbolic notation, to express mathematical ideas with precision. Representing, discussing, justifying, conjecturing, reading, writing, presenting, and listening to mathematics will help students clarify their thinking and deepen their understanding of the mathematics being studied. Mathematical communication becomes visible where learning involves participation in mathematical discussions.

Mathematical Reasoning

Students will recognize reasoning and proof as fundamental aspects of mathematics. Students will learn and apply inductive and deductive reasoning skills to make, test, and evaluate mathematical statements and to justify steps in mathematical procedures. Students will use logical reasoning to analyze an argument and to determine whether conclusions are valid. In addition, students will use number sense to apply proportional and spatial reasoning and to reason from a variety of representations.

Mathematical Connections

Students will build upon prior knowledge to relate concepts and procedures from different topics within mathematics and see mathematics as an integrated field of study. Through the practical application of content and process skills, students will make connections among different areas of mathematics and between mathematics and other disciplines, and to real-world contexts. Science and mathematics teachers and curriculum writers are encouraged to develop mathematics and science curricula that support, apply, and reinforce each other.

Mathematical Representations

Students will represent and describe mathematical ideas, generalizations, and relationships using a variety of methods. Students will understand that representations of mathematical ideas are an essential part of learning, doing, and communicating mathematics. Students should make connections among different representations ~~—~~ physical, visual, symbolic, verbal, and contextual ~~—~~ and recognize that representation is both a process and a product.

Instructional Technology

The use of appropriate technology and the interpretation of the results from applying technology tools must be an integral part of teaching, learning, and assessment. However, facility in the use of technology shall not be regarded as a substitute for a student's understanding of quantitative and algebraic concepts and relationships or for proficiency in basic computations. Students must learn to use a variety of methods and tools to compute, including paper and pencil, mental arithmetic, estimation, and calculators. In addition, graphing utilities, spreadsheets, calculators, dynamic applications, and other technological tools are now standard for mathematical problem solving and application in science, engineering, business and industry, government, and practical affairs.

Calculators and graphing utilities should be used by students for exploring and visualizing number patterns and mathematical relationships, facilitating reasoning and problem solving, and verifying solutions. However, according to the National Council of Teachers of Mathematics, "… the use of calculators does not supplant the need for students to develop proficiency with efficient, accurate methods of mental and pencil-and-paper calculation and in making reasonable estimations." State and local assessments may restrict the use of calculators in measuring specific student objectives that focus on number sense and computation. On the grade three state assessment, all objectives are assessed without the use of a calculator. On the state assessments for grades four through seven, objectives that are assessed without the use of a calculator are indicated with an asterisk (*).

Computational Fluency

Mathematics instruction must develop students' conceptual understanding, computational fluency, and problem-solving skills. The development of related conceptual understanding and computational skills should be balanced and intertwined, each supporting the other and reinforcing learning.

Computational fluency refers to having flexible, efficient and accurate methods for computing. Students exhibit computational fluency when they demonstrate strategic thinking and flexibility in the computational methods they choose, understand and can explain, and produce accurate answers efficiently.

The computational methods used by a student should be based on the mathematical ideas that the student understands, including the structure of the base-ten number system, number relationships, meaning of operations, and properties. Computational fluency with whole numbers is a goal of mathematics instruction in the elementary grades. Students should be fluent with the basic number combinations for addition and subtraction to 20 by the end of grade two and those for multiplication and division by the end of grade four. Students should be encouraged to use computational methods and tools that are appropriate for the context and purpose.

Algebra Readiness

The successful mastery of Algebra I is widely considered to be the gatekeeper to success in the study of upper-level mathematics. "Algebra readiness" describes the mastery of, and the ability to apply, the *Mathematics Standards of Learning*, including the Mathematical Process Goals for Students, for kindergarten through grade eight. The study of algebraic thinking begins in kindergarten and is progressively formalized prior to the study of the algebraic content found in the *Algebra I Standards of Learning*. Included in the progression of algebraic content is patterning, generalization of arithmetic concepts, proportional reasoning, and representing mathematical relationships using tables, symbols, and graphs. The *K-8 Mathematics Standards of Learning* form a progression of content knowledge and develop the reasoning necessary to be well-prepared for mathematics courses beyond Algebra I, including Geometry and Statistics.

Equity

> **"**Addressing equity and access includes both ensuring that all students attain mathematics proficiency and increasing the numbers of students from all racial, ethnic, linguistic, gender, and socioeconomic groups who attain the highest levels of mathematics achievement."
>
> – National Council of Teachers of Mathematics

Mathematics programs should have an expectation of equity by providing all students access to quality mathematics instruction and offerings that are responsive to and respectful of students' prior experiences, talents, interests, and cultural perspectives. Successful mathematics programs challenge students to maximize their academic potential and provide consistent monitoring, support, and encouragement to ensure success for all. Individual students should be encouraged to choose mathematical programs of study that challenge, enhance, and extend their mathematical knowledge and future opportunities.

Student engagement is an essential component of equity in mathematics teaching and learning. Mathematics instructional strategies that require students to think critically, to reason, to develop problem-solving strategies, to communicate mathematically, and to use multiple representations engages students both mentally and physically. Student engagement increases with mathematical tasks that employ the use of relevant, applied contexts and provide an appropriate level of cognitive challenge. All students, including students with disabilities, gifted learners, and English language learners deserve high-quality mathematics instruction that addresses individual learning needs, maximizing the opportunity to learn.

Data ~~and Society~~in Context - Understanding data science facilitates critical examination of questions ~~in different parts of society~~ and supports informed data-driven decision~~-~~ making.

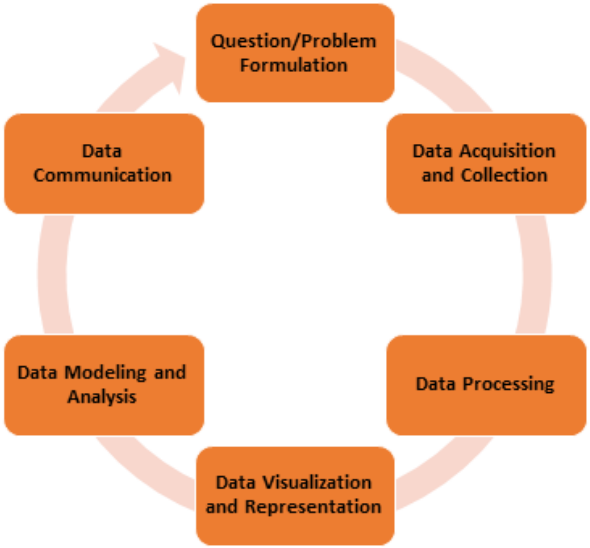
DS.1† **The student will identify specific examples of ~~societal~~ real-world problems that can be effectively addressed using data science.**

Understanding the Standard	Essential Knowledge and Skills
● There are characteristics of problems in ~~society~~ the real-world that best lend themselves to be analyzed using the data cycle. ● Solutions addressed by Data Science include conjectures that can be supported or refuted by measurements or observations. ● The iterative stages of the data cycle include: – Question/Problem Formulation - Identify the driving question for the problem being solved – Data Acquisition & Collection - Collect and clean data to assist with multiple ways to solve a problem – Data Processing - Manipulate data to make it usable through a predetermined process – Data Visualization & Representation - Connect visual representations to brainstorm solutions – Data Modeling & Analysis - Build a prototype of a model, test, and iterate – Data Communication - Effectively communicate data driven solution based on context and audience	**The student will use problem solving, mathematical communication, mathematical reasoning, connections, and representations to** ● Identify and explain characteristics that best lend themselves to a data driven approach to problem solving. ● Formulate questions based on context. ● Understand the type of data relevant to the context of the question at hand. ● Define relationships between variables and constant relationships. ● Create a hypothesis of interest in terms of measurable data. ● Define the stages of the data cycle and how each stage is related to the other. ● Identify and explain constraints of the data-driven approach.

Data ~~and Society~~in Context - Understanding data science facilitates critical examination of questions ~~in different parts of society~~ and supports informed data-driven decision~~-~~ making.

DS.1† **The student will identify specific examples of ~~societal~~ real-world problems that can be effectively addressed using data science.**

Data Cycle



- The data science cycle is an iterative process.

† Standard should be included in a one-semester course in Data Science.

Data ~~and Society~~in Context - Understanding data science facilitates critical examination of questions ~~in different parts of society~~ and supports informed data-driven decision~~-~~ making.

DS.2 The student will be able to formulate a top-down plan for data collection and analysis, with quantifiable results, based on the context of a problem.

Understanding the Standard	Essential Knowledge and Skills
• A data project plan ensures effective communication and agreement at all phases of the data science project. • A data project plan allows effective execution on time and under budget. • A data project plan allows us to understand the tools, resources and architecture needed to ensure a successful project. • Project deliverables are the things you create to help you fulfill the objective while KPI stands for key performance indicator, a quantifiable measure of success of the project as a whole. • Sampling bias in the data collection process include, but are not limited to, confirmation, selection, and outliers. • ~~—~~Sampling must be purposeful to infer trends and characteristics in the data being collected. • Nonrandom sampling techniques, such as convenience, quota, judgment, and snowball, may result in a non-representative sample that does not produce generalizable results.	**The student will use problem solving, mathematical communication, mathematical reasoning, connections, and representations to** • Design a data project plan, which is aligned with the data science cycle, that includes the following components: – Definition of the goal of the project as it pertains to a ~~societal~~ real-world problem; – Identification of the various parameters of the problem and stakeholders; – A timeline for the project with deliverables; – Key Performance Indicators (KPI) for the successful data project deliverables; – Resource needs and tools for the project; – ~~Ethical~~ Bias considerations for the sampling process of ~~around~~ the project; and – Limitations of the project. • Given the context and parameters of a problem, choose from among various sampling techniques, which may include ○ simple random; ○ systematic; ○ stratified; ○ cluster; to justify the sampling methodology of the project design and implementation.

Data ~~and Ethics~~ Bias ~~-~~ ~~Ethical implications~~Data bias may result from the types of methods used for data collection, processing, representation, analysis, and use.

DS.3† **The student will recognize the importance of data literacy ~~in global citizenship~~ and develop an awareness of how the analysis of data can be used in problem solving to affect ~~positive changes and mitigate negative consequences~~ change and create innovative solutions.**

Understanding the Standard	Essential Knowledge and Skills
• Data literacy is the ability to read data, work with data and communicate about data by putting it in proper context and asking relevant/clarifying questions to determine/identify data bias. • Data literacy helps to recognize, sort and filter through data biases that leads to improved decision ~~-~~making in data collection and reporting. • Data privacy and consumer protection are important issues that affect individuals and organizations. • Historical instances of government and private data breaches provide examples of the considerations of privacy in data. • Data bias occurs when data does not include variables that properly capture the phenomenon we want to predict.	**The student will use problem solving, mathematical communication, mathematical reasoning, connections, and representations to** • Formulate relevant/clarifying questions to identify potential data biases ~~in data~~ presented in existing analyses/visualizations. • Effectively read data summaries and visualizations and explain/translate into non~~-~~technical terms in proper context. • Identify potential data biases in terms of data presented and discuss the potential effects of such biases in terms of how they could affect data analysis and decision~~-~~ making. • Identify privacy and consumer protection issues that might be a result of how data is presented. • Describe the types of data that business, industry, and government entities collect ~~about people~~ and possible ways the data is used.

† Standard should be included in a one-semester course in Data Science.

Data ~~and Ethics~~ Bias ~~-~~ ~~Ethical implications~~Data bias may result from the types of methods used for data collection, processing, representation, analysis, and use.

DS.4 The student will be able to identify data biases in the data collection process, and understand the ~~basic ethical~~ implications and privacy issues surrounding data collection and processing.

Understanding the Standard	Essential Knowledge and Skills
• ~~Ethical i~~Various implications can result from the types of data collection methods used. • Privacy and consumer protection are considerations when data ~~is~~ are collected. • ~~Different types of biases can occur while collecting data, including implicit and explicit biases.~~ • There are producers, publishers, consumers and decision ~~-~~makers of data. – Producer of data: data are~~is~~ obtained through some source~~-~~ open source, sensor equipment, third party organization/source, external source – Publisher of data: entity that acquires, manages, stores, makes available the data – Consumer of data: develops products/applications to support the decision~~-~~ making – Decision maker of data: uses the products/applications to make decisions	**The student will use problem solving, mathematical communication, mathematical reasoning, connections, and representations to** • Identify data biases in the data collection process that include, but are not limited to, confirmation, selection, outliers, overfitting / ~~underfitting~~under fitting, and confounding and describe mitigation strategies for these biases. • Provide examples of sampling biases in terms of data collection and the potential effects ~~of such biases~~. • Identify and describe data biases ~~apparent in the data given the biases of the~~as a producer as well as ~~those of the~~a consumer/decision maker of ~~the~~ data. • ~~Identify potential sources of bias, given a specific data collection scenario.~~ • Describe how the data collection process should be focused, relevant, and limited to the scope of the data project plan. • Describe ~~basic ethical/~~privacy ~~issues~~ considerations~~possible~~ in the collection of data as both a consumer and producer.

Data and Communication - Data visualizations are used to communicate insights about complex data sets to support ~~an audience in~~ making decisions.

DS.5[†] **The student will use storytelling as a strategy to effectively communicate with data.**

Understanding the Standard	Essential Knowledge and Skills
• Storytelling with data involves combining context, visualizations and a narrative to communicate the idea behind a data science project effectively. Narrative, which is the crux of storytelling, is the way we simplify and make sense of complex data by supplying context, insight, and interpretation to make the analysis more applicable and relevant. • Communicating with data using storytelling involves concrete steps: – Understanding context, – Selecting a visual, – Eliminating clutter, – Focus attention, and – Telling a story. • Data storytelling requires accuracy in presenting information and critical thinking in ~~consuming~~ consuming ~~done incorrectly can lead to incomplete or misleading~~ information ~~and~~ to make conclusions.	**The student will use problem solving, mathematical communication, mathematical reasoning, connections, and representations to** • Define storytelling and explain the importance of storytelling as a strategy to communicate the idea behind and results of a data science project effectively. • Explain the steps involved in data storytelling and how it relates to the data cycle. – Effectively identify a story worth telling based on the data (looking for trends, correlations, outliers) and by asking a question or forming a hypothesis based on insight and audience. – Effectively selecting visualizations that simplify the information, highlight the most important data, and communicate key points quickly. – Effectively simplifying the information presented to make it more concise and focusing the audience's attention on the key parameters that support the student's hypothesis. – Effectively form a narrative based on data available to provide context, insight, interpretation to make the analysis more relevant to a given audience. • Explain how data storytelling ~~done incorrectly can lead to incomplete~~should include complete and ~~accurate or misleading information~~accurate information, ~~and conclusions by eliminating/substituting, manipulating, cherry picking data, or creating in~~ ~~and~~consistent and consistent visuals ~~or ineffective narratives~~for effective communication.

[†] Standard should be included in a one-semester course in Data Science.

Data and Communication - Data visualizations are used to communicate insights about complex data sets to support ~~an audience in~~ making decisions.

DS.6† **The student will justify the design, use and effectiveness of different forms of data visualizations.**

Understanding the Standard	Essential Knowledge and Skills
• The goal of data visualization is to distill large datasets into visual graphics to allow for easy understanding of complex relationships within data. • Computer-based visualization systems provide visual representations of data sets designed to help end users to carry out tasks more effectively. Data visualization includes analysis, design, and construction. • Task questions may include: What questions does the user want to answer? What problem is to be solved? Which decisions is the user trying to make? What outcomes are desired? What story does the user want to tell? What tasks should the user perform? • Choosing a visualization based on data type and the message communicated reveals trends so the audience can easily understand the significance of the findings from the data set. • Data set types in visualizations include but are not limited to: tabular; network; spatial; and textual. Tabular data may be represented in two-dimensional (row by column) or multidimensional tables. Networks may include nodes and links and trees. Spatial data sets may be categorized as continuous fields as in grids of position and geometric such as in maps. • Inputs for visualizations include data set types and tasks. Data attributes may be categorical, ordinal or quantitative with special cases for time and space. • Data visualizations may include both conventional and emerging types based on function in the context of the data.	**The student will use problem solving, mathematical communication, mathematical reasoning, connections, and representations to** • Conduct exploratory data analysis using visualization. • Formulate questions from exploration of a data set to consider how data will communicate a story. • Determine the effectiveness of different data visualization choices based on the data context from conventional statistical charts to unconventional/emerging data visualizations to more complex visualizations. • Create a visualization of a data set and summarize the representation using the context of the data. • Compare two or more different representations to ensure the design communicates the features and behavior of data sets. • Justify design choices (based on data set type, size, context and audience) of data visualizations to highlight important features, trends, and insights.

Data and Communication - Data visualizations are used to communicate insights about complex data sets to support ~~an audience in~~ making decisions.

DS.6† **The student will justify the design, use and effectiveness of different forms of data visualizations.**

- Data insights from visualizations can be shared in different ways including: live or virtual presentations; dashboards; embedded into applications; and/or broadcast to audiences through data-driven alerts or communications.
- The choice of a suitable technological tool allows students to create and compare multiple visualizations of the same data set.
- Connections can be made among summary information from statistical analysis to visualizations of the same data set.
- Numerous forms of data visualizations exist and are often chosen based on the intended function of the visualization.

Chart Selection for Data Visualization by Function

	Comparisons	Proportions	Relationships	Hierarchy	Location	Distribution	Patterns	Range	Data Over Time	Analyzing Text	Movement/ Flow	Financial	Uncertainty/ Error
Area Graph/Plot Stacked Area Graph/Plot	X						X		X				
Area Bands													X
Bar Graph Stacked Bar Graph	X	X					X						
Box and Whisker Plot	X					X	X	X					
Bubble Chart/-Map	X	X	X		X	X	X		X				
Candlestick Chart								X	X			X	
Chord Diagram			X										
Choropleth Map					X								
Circle Packing		X		X									
Confidence Strips													X
Connections Map			X			X					X		
Data Over Geographical Region					X								
Density Chart/Plot						X	X						
Donut Chart		X											
Dot Map					X	X	X						
Dot Matrix		X				X							
Error Bars													X
Flow Map					X	X					X		

Data and Communication - Data visualizations are used to communicate insights about complex data sets to support ~~an audience in~~ making decisions.

DS.6† **The student will justify the design, use and effectiveness of different forms of data visualizations.**

Gantt Chart							X	X					
Heat Map			X						X				
Histogram	X					X	X	X	X				
Kagi Chart												X	
Line Graph	X					X	X		X				
Marimekko Chart			X										
Multivariable Bar Chart						X	X						
Parallel Sets											X		
Pie Chart		X											
Population Pyramid						X	X						
Renko Chart												X	
Sankey Diagram											X		
Scatterplot			X			X	X						
Span Chart								X					
Spiral Plot									X				
Stream Graph									X				
Sunburst				X									
Tree Diagram/Map		X	X	X									
Two-Way Tables	X												
Venn Diagram			X										
Violin Chart								X					
Waterfall Chart												X	
Word Cloud		X								X			

Data Modeling ~~–~~ Mathematical models are used to predict future, unobserved data values.

DS.7 The student will be able to assess reliability ~~and validity~~ of source data in preparation for mathematical modeling.

Understanding the Standard	Essential Knowledge and Skills
● Understanding the characteristics of a reliable data source will allow for more effective analysis. ● ~~Understanding the difference between a reliable and a valid data source compared to statistical validity and reliability in research analysis is important.~~ ● There are different aspects of data reliability: – Data can be considered ~~V~~validity ~~– is the data correctly when~~valid when it is formatted and stored ~~in the right way?~~in a consistent structure; – Data is ~~C~~completeness ~~– does the dataset when~~complete when it includes all values ~~for all the fields required by your system?~~required by the context; and – Data is ~~u~~Uniqueness ~~– is the data if~~unique if it is free from duplicates and ~~extraneous~~dummy ~~entries~~extraneous entries~~?~~. ● Data validation or input validation is a method for checking the accuracy and quality of source data, typically performed prior to importing and processing so that data analysis results are accurate.	**The student will use problem solving, mathematical communication, mathematical reasoning, connections, and representations to** ● Explain why determining the reliability ~~and validity~~ of big data sources is a key skill that data scientists use to build data trust across an organization. ● ~~Assess the validity of different data sources.~~ Describe the difference between reliability of a data source compared to statistical reliability and validity in research analysis. ● Assess processing source data for reliability based on validity, completeness and uniqueness.

Data Modeling ~~–~~ – Mathematical models are used to predict future, unobserved data values.

DS.8† **The student will be able to acquire and prepare big data sets for modeling and analysis.**

Understanding the Standard	Essential Knowledge and Skills
• Data can be collected or acquired from reliable existing data sources. • The purpose of sampling is to provide sufficient information so that population characteristics may be inferred. • Data preparation ~~helps catch~~supports identifying errors before processing. • Cleaning and reformatting data sets ensures that all data used in analysis will be high quality. • Higher quality data can be processed and analyzed more quickly and efficiently. • The process involved in preparing the data set for modeling and analysis involves one or more of the following sub-steps: – Ingest~~ing~~/~~wrangling~~ wrangle the data~~,~~, which includes: ○ Sort (arrange) - order rows by the value or characters of a variable, or a selection of them; ○ Select - choose columns in a dataset based on a defined criteria; ○ Filter - remove parts of rows of a dataset during analysis; ○ Replace - convert specific characters (e.g., convert numerical characters to data and time formats) or re-code variables to fit models. – Clean~~ing~~ the data~~,~~; – ~~Formatting~~ Format and ~~E~~enrich~~ing~~ the data~~,~~; and – ~~Combining~~ Combine and ~~storing~~ store the data.	**The student will use problem solving, mathematical communication, mathematical reasoning, connections, and representations to** • Explain the pros and cons of collecting data vs. acquiring it from existing sources. • Utilize tools and their functions to combine and store data by: – Removing data that ~~is~~are incomplete, incorrect or duplicated; – Removing extraneous data or outliers; and – Standardizing data to conform to contextual norms (e.g., privacy, sensitive data). [Moved from DS.10] • ~~Explain various sampling techniques and their effectiveness - probability (simple random, systematic, stratified, and cluster sampling methods) vs. non-probability (convenience, quota, judgment, snowball).~~ [Reworded and Moved to DS.2] • Utilize tools and their functions to clean and validate data by: – Merging multiple data sets for efficiency purposes; and – Optimizing storage of data based on volume, velocity, and variety. [Moved from DS.10] • Apply matrix operations using algebraic methods (with the support of technology tools) to: – Wrangle the data (sort, select, filter, and replace); – Clean the data; – Format and enrich the data; and – Combine and store the data. • Read data from different sources for preparation and analysis. • Identify important parameters about a big data set based on the context of data collected/acquired.

† Standard should be included in a one-semester course in Data Science.

Data Modeling ~~–~~ Mathematical models are used to predict future, unobserved data values.

DS.9† **The student will select and analyze data models to make predictions, while assessing accuracy and sources of uncertainty.**

Understanding the Standard	Essential Knowledge and Skills
• Data prediction involves extrapolating the data beyond the current data set and providing confidence values for those estimates. • It is important to be able to distinguish between the "noise" in the data and relevant data. Every measurement is composed of true value, bias and random noise. This noise is the source of uncertainty. • Mathematical models will be used to make data predictions based on the behavior of the data. • Data prediction ~~is~~ may be limited by the ~~fact that it is based on an~~ assumption that historical patterns are a good predictor of future outcomes. • Overfitting the data can lead to inaccurate results. • ~~Ethical c~~Considerations based on data bias need to be taken into account during feature selection when trying to predict future outcomes. • ~~Students will understand t~~The fundamentals of numerical methods, allow for ~~especially their~~ further understanding of the application, limitations, and ~~potentials.~~pitfalls of the model.	**The student will use problem solving, mathematical communication, mathematical reasoning, connections, and representations to** • Identify factors that contribute to the overall behavior of a data set, including true values, bias and noise. • Fit models based on the behavior of the data, including models of univariate and bivariate data, in order to make predictions. • Distinguish between linear and non-linear associations between variables using visualizations. • Identify models that are overly complex and therefore fitting to random noise which decreases their predictive accuracy. • Use regression techniques to ~~P~~perform ~~feature~~ selection of ~~optimal~~ optimal ~~to choose~~ features. • ~~which are relevant to study while r~~Recognizinge the potential ~~ethical~~ implications of removing features. • Select the ~~best~~ optimal model for a data set from among a large collection of models, using technological tools.

† Standard should be included in a one-semester course in Data Science.

Data Modeling ~~–~~ Mathematical models are used to predict future, unobserved data values.

DS.10† **The student will be able to summarize and interpret data represented in both conventional and emerging visualizations.**

Understanding the Standard	Essential Knowledge and Skills
• Characteristics of data sets can be summarized graphically by using visual representations of the distribution and numerically with measures of central tendency and measures of variation or dispersion. • Descriptive statistics summarize the characteristics of a data set. ~~Inferential statistics allow you to test a hypothesis or generalize findings from a sample group to a larger population.~~ • Statistical summaries have the potential to lose information. Representing all the data through visualizations is important to confirm expected patterns, find unexpected patterns, and to assess the validity of the selected statistical model. ~~• Define emerging visualizations and describe summarization of characteristics and relationships among variables including:~~ ~~– A heat map uses color to show changes and magnitude of a third variable to a two-dimensional plot.~~ ~~– A bubble chart is a multivariable graph that is a cross between a scatterplot and a proportional area chart. Each plotted point then represents a third variable by the area of its circle.~~ • Visualizations are a key to validating underlying assumptions such as data being normally distributed and having no correlation between independent variables.	**The student will use problem solving, mathematical communication, mathematical reasoning, connections, and representations to** • Apply descriptive statistics to explain measures of central tendency and measures of variability/dispersion to describe center and spread in visualizations of distributions. • Define emerging visualizations and describe summarization of characteristics and relationships based on audience and purpose which may include: – A heat map, which uses color to show changes and magnitude of a third variable to a two-dimensional plot. – A bubble chart, which is a multivariate graph that is both a scatterplot and a proportional area chart. Typically, each plotted point then represents a third variable by the area of its circle.~~.~~ • Interpret various emerging visualizations by describing patterns, trends and relationships between and among the variables.

- Selected Charts for Data Visualization based on types and number of variables:

	Univariate	**Bivariate**	**Three Variables of Higher**
Quantitative	Dotplots Stemplots Histograms Box and Whisker Plots	Scatterplots Line Plots 2-D Histograms	3-D Scatterplot 3-D Lineplot Heat Map Bubble Chart
Categorical	Bar Charts Pie Charts	Two-Way Tables Segmented Bar Graphs	Multivariate Bar Graphs

† Standard should be included in a one-semester course in Data Science.

Data Modeling ~~–~~ Mathematical models are used to predict future, unobserved data values.

DS.11 The student will ~~use hypothesis formulation~~select statistical models and use goodness of fit testing to extract actionable knowledge directly from data.

Understanding the Standard	Essential Knowledge and Skills
• There are key differences between observed and theoretical probabilities. • The different types of distribution of data vary according to the context and are important to predict future outcomes • While causation and correlation can exist at the same time, correlation does not imply causation. • Categorical variables can also be analyzed using specific tests. • ~~Unsupervised learning~~Technology tools can be used ~~to~~ ~~includes identify~~to identify~~ing~~ meaningful clusters of data and associated sets of data points. Methods like clustering can be used to identify meaningful relationships between data observations in the form of similarities. When visualizing clustering methods, these similarities show up as "closeness" between plotted data points or the tendency of similar points to group together. • It is important to have a toolbox of different statistical models for modeling a variety of phenomena (Binomial, Poisson, exponential, etc.) • Histogram comparisons, Chi-squared tests, and other methods are used to test goodness of fit.	**The student will use problem solving, mathematical communication, mathematical reasoning, connections, and representations to** • Calculate the theoretical probability of random events and compare them to the observed frequencies. • Describe the normal curve determined by the mean and standard deviation of a univariate data set. • Fit non-linear models to data sets and use these models to predict unobserved data values. • Select pairs of variables that identify meaningful clusters of data. • Select an appropriate statistical distribution and test its goodness of fit based on the context of the data being analyzed. Statistical distributions may include, but are not limited to ○ ~~including the n~~Normal ~~distribution;~~ ○ ~~Binonomial~~Bionomical; and ○ Poisson. ~~and other discrete and continuous distributions.~~

Data and Computing – Technology is used to effectively prepare, analyze, and communicate with data.

DS.12† **The student will be able to select and utilize appropriate technological tools and functions within those tools to process and prepare data for analysis.**

Understanding the Standard	Essential Knowledge and Skills
• Data can be imported, processed, and exported (if necessary) using technology tools. • Organizing data using technology tools aids in exploration. • Technology tools can be used to handle address missing entries, errors, or duplicates in the data. • The process of decision- making that occurs during the importing or extracting, processing, cleaning and formatting of data uses a choice of tools: technological applications, coding, and web. • It is important that this data pre-processingThe technology procedure for data pre-processing is clearly explained and documented for future replication and decision- making.	**The student will use problem solving, mathematical communication, mathematical reasoning, connections, and representations to** • Utilize technology tools to be able to access data effectively from multiple sources (e.g., tables, column separated values, spreadsheets, documents, databases). • Utilize tools and functions (in tools) to effectively explore the data for issues and errors before beginning to process it. • Define the (tools and technological) process to optimally ingest data and to export data after processing. • Utilize tools to format and store the data appropriately to allow for effective analysis. • Utilize tools and functions (in tools) to clean and validate data by: – Removing data that is incomplete, incorrect or duplicated; – Removing extraneous data or outliers; and – Standardizing data to conform to contextual norms (e.g., privacy, sensitive data). [Moved to DS.8] • Utilize tools and their functions to Ccombine and store data by: – Merging multiple data sets for efficiency purposes; and – Optimizing the storage of data based on volume, velocity and variety. [Moved to DS.8] • Define and document the process of ingesting, formatting and cleaning data for future decision making by: – Making data more easily understood by a wider audience; and – Connecting data with existing contextual data.

† Standard should be included in a one-semester course in Data Science.

Data and Computing – Technology is used to effectively prepare, analyze, and communicate with data.

DS.13† **The student will be able to select and utilize appropriate technological tools and functions within those tools to analyze and communicate data effectively.**

Understanding the Standard	Essential Knowledge and Skills
• Certain technological tools can be used to generate conventional and unconventional visualizations of data to explore patterns and/or analyze a large data set. • Various technological tools have pre-built mathematical and statistical functions that allow for efficient exploration and analysis. • Coding tools can allow for effective storage and extraction of data for more efficient analysis. • Some technological tools have other functions that are useful to organize, summarize and gain insight from data. • Visualization tools offer a variety of conventional and unconventional visualizations to help communicate our ideas to a wide audience.	**The student will use problem solving, mathematical communication, mathematical reasoning, connections, and representations to** • Select and utilize technology tools to effectively generate conventional and unconventional visualizations of data to explore patterns and/or analyze a large data set. • Utilize specific functions in technology tools to perform descriptive and inferential statistical analysis. • Utilize coding to store and extract data more effectively for data analysis. • Select and apply features of technology tools effectively to organize, summarize and gain insight from data. • Select the appropriate visualization based on context and audience and create it using technology tools to effectively communicate an idea.

† Standard should be included in a one-semester course in Data Science.

Mathematics Standards of Learning - Data Science (Adopted March, 2022)

The following standards outline the content of a one-year course in Data Science. If a one-semester course is desired, the standards with a dagger (†) would apply. The *Data Science Standards of Learning* provide an introduction to the learning principles associated with analyzing big data.

Through the use of open source technology tools, students will identify and explore problems that involve the use of relational database concepts and data-intensive computing to find solutions and make generalizations. Students will engage in a data science problem-solving structure to interact with large data sets as a means to formulate problems, collect and clean data, visualize data, model using data, and communicate effectively about data formulated solutions.

Data ~~and~~ in Context ~~Society~~ - Understanding data science facilitates critical examination of questions ~~in different parts of society~~ and supports informed data-driven decision-making.

DS.1† The student will identify specific examples of ~~societal~~ real-world problems that can be effectively addressed using data science.

DS.2 The student will be able to formulate a top down plan for data collection and analysis, with quantifiable results, based on the context of a problem.

Data ~~and Ethics~~ Bias - ~~Ethical implications~~Data bias may result from the types of methods used for data collection, processing, representation, analysis, and use.

DS.3† The student will recognize the importance of data literacy ~~in global citizenship~~ and develop an awareness of how the analysis of data can be used in problem solving to affect ~~positive changes and mitigate negative consequences~~change and create innovative solutions.

DS.4 The student will be able to identify data biases in the data collection process, and understand the ~~basic ethical~~ implications and privacy issues surrounding data collection and processing.

Data and Communication - Data visualizations are used to communicate insights about complex data sets to support ~~an audience in~~ making decisions.

DS.5† The student will use storytelling as a strategy to effectively communicate with data.

DS.6† The student will justify the design, use, and effectiveness of different forms of data visualizations.

Data Modeling - Mathematical models are used to predict future, unobserved data values.

DS.7 The student will be able to assess reliability ~~and validity~~ of source data in preparation for mathematical modeling.

DS.8† The student will be able to acquire and prepare big data sets for modeling and analysis.

DS.9† The student will select and analyze data models to make predictions, while assessing accuracy and sources of uncertainty.

DS.10† The student will be able to summarize and interpret data represented in both conventional and emerging visualizations.

DS.11 The student will ~~use hypothesis formulation~~select statistical models and use goodness of fit testing to extract actionable knowledge directly from data.

Data and Computing ~~–~~ Technology is used to effectively prepare, analyze, and communicate with data.

DS.12† The student will be able to select and utilize appropriate technological tools and functions within those tools to process and prepare data for analysis.

DS.13† The student will be able to select and utilize appropriate technological tools and functions within those tools to analyze and communicate data effectively.

Attachment C

Data Science Standards of Learning and *Data Science Standards of Learning Curriculum Framework Anticipated Implementation Timeline and Communication Plan*

Date	Action	Communication	Method
March 2022	The Department of Education seeks school divisions to submit applications to participate in a pilot implementation of the *Data Science Standards of Learning* and *Data Science Standards of Learning Curriculum Framework* during the 2022-2023 school year.	Announce availability of an application process for school divisions to participate in a pilot implementation of the *Data Science Standards of Learning* and *Data Science Standards of Learning Curriculum Framework* during the 2022-2023 school year.	Superintendent's Email
April 2022	The Department of Education will notify school divisions selected to pilot implementation of the *Data Science Standards of Learning* and *Data Science Standards of Learning Curriculum Framework.*	Directly contact school division representatives regarding selection to participate in the pilot implementation of the *Data Science Standards of Learning* and *Data Science Standards of Learning Curriculum Framework.*	Direct Email to selected school divisions from VDOE Staff
May 2022	Final versions of the *Data Science Standard of Learning* and *Data Science Standards of Learning Curriculum Framework* are posted on the VDOE website	Announce posting of final versions of the *Data Science Standards of Learning* and *Data Science Standards of Learning Curriculum Framework.*	Superintendent's Memo, VDOE social media, TeacherDirect, and other communication channels
Summer 2022	VDOE provides professional development to a cohort of teachers based on division participation in a pilot implementation	Cohort teachers and their division contacts will be provided professional development information	Direct Email to selected school divisions from VDOE Staff
August 2022-June 2023	School divisions selected for pilot implementation offer Data Science courses Department of Education and select teachers from the pilot implementation cohort conducts ongoing review of draft *Data Science instructional resources*	Select communication with school divisions selected for pilot implementation	Email communication and on-site visitations by Department of Education staff
Summer 2023	VDOE provides professional development to all school divisions wishing to offer a Data Science Course during the 2023-2024 school year	Announce professional development	Superintendent's Memo, VDOE social media, TeacherDirect, and other communication channels

Attachment C

Data Science Standards of Learning and *Data Science Standards of Learning Curriculum Framework Anticipated Implementation Timeline and Communication Plan*

Date	Action	Communication	Method
2023-2024 School Year	Full implementation of the *Data Science Standards of Learning* and *Data Science Standards of Learning Curriculum Framework*		