Data Science Honors Essential Curriculum

The Mathematical Practices

The Standards for Mathematical Practice describe varieties of expertise that mathematics educators at all levels should seek to develop in their students.

1. Make sense of problems and persevere in solving them
2. Reason abstractly and quantitatively
3. Construct viable arguments and critique the reasoning of others
4. Model with mathematics
5. Use appropriate tools strategically
6. Attend to precision
7. Look for and make use of structure
8. Look for and express regularity in repeated reasoning

The Mathematical Content Standards

The Student Learning Outcomes (Essential Curriculum) that follow identify the knowledge, skills, and mathematical understandings students are expected to develop through instruction and learning in Howard County Public Schools. The outcomes are designed to promote a balanced combination of procedure and understanding, with opportunities for students to develop and apply the Mathematical Practices alongside the content.

Students with a strong understanding of mathematics are better positioned to make connections, represent problems in multiple ways, communicate and justify their thinking, apply mathematics to practical situations, and adapt their strategies when familiar procedures are not efficient. In short, conceptual understanding provided the foundation for students to engage meaningfully in the Mathematical Practices.

Unit 1: Data Tells a Story

Student Learning Outcomes:

- Conduct data investigations as an iterative process, including: formulating statistical questions, exploring and analyzing data, revising assumptions or approaches, and communicating evidence-based conclusions.
- Use exploratory data analysis to identify patterns, test initial ideas, and determine next steps in a data investigation.
- Evaluate the quality, credibility, and relevance of data sources for a given investigation.
- Analyze data collection processes, including sampling methods, representativeness, potential bias, reliability, and ethical considerations such as privacy, participant harm, and informed consent.
- Clean and prepare real-world datasets by addressing missing values, duplicates, inconsistent formats, and incomplete or unstructured data.

- Organize and transform data to support analysis, including joining datasets, restructuring variables, and converting data into usable formats.
- Create and interpret purpose-focused data visualizations, and select appropriate representations for univariate, bivariate, and multivariate data.
- Evaluate and critique data visualizations for clarity, accuracy, accessibility, and potential sources of misinterpretation.
- Communicate insights from data, using visual representations to construct and support data-informed arguments that explain findings, acknowledge limitations, and address uncertainty.
- Identify and analyze sources of bias in datasets, models, and algorithms, and explain how these biases may affect conclusions or decisions.
- Evaluate ethical considerations in data practices (e.g., privacy, consent, transparency, and responsible data stewardship).
- Assess the societal impacts of data-driven systems, and discuss how data science can both reflect and shape social outcomes.
- Make reasoned judgments about responsible data use, considering, for example, fairness, accountability, and the limitations of data and algorithms.

Unit 2: The Data of our Community: Learning from Data Distributions

Student Learning Outcomes:

- Describe and interpret variability and distributions in data using graphical displays and summary statistics, including measures of center and spread, consideration of outliers, and how distribution shape and context inform conclusions.
- Use probability and simulation to reason about randomness, estimate the likelihood of events, and interpret probabilistic outcomes in context.
- Analyze relationships between variables using visualizations and statistical measures to describe patterns, direction, and strength of association, including interpreting correlation and R2 to evaluate how well a model explains variability in context.
- Develop and evaluate statistical models to represent relationships in data, including selecting and interpreting appropriate function types (e.g., linear, exponential, or other nonlinear models) and recognizing the usefulness and limitations of models in representing real-world phenomena.

Unit 3: Water in Your Life: Bivariate Data and Causality v. Spurious Correlation

Student Learning Outcomes:

- Conduct data investigations as an iterative process, including: formulating statistical questions, exploring and analyzing data, revising assumptions or approaches, and communicating evidence-based conclusions.

- Use exploratory data analysis to identify patterns, test initial ideas, and determine next steps in a data investigation.
- Describe and interpret variability and distributions in data using graphical displays and summary statistics, including measures of center and spread, consideration of outliers, and how distribution shape and context inform conclusions.
- Use probability and simulation to reason about randomness, estimate the likelihood of events, and interpret probabilistic outcomes in context.
- Analyze relationships between variables using visualizations and statistical measures to describe patterns, direction, and strength of association, including interpreting correlation and R2 to evaluate how well a model explains variability in context.
- Develop and evaluate statistical models to represent relationships in data, including selecting and interpreting appropriate function types (e.g., linear, exponential, or other nonlinear models) and recognizing the usefulness and limitations of models in representing real-world phenomena.
- Create and interpret purpose-focused data visualizations, and select appropriate representations for univariate, bivariate, and multivariate data.
- Evaluate and critique data visualizations for clarity, accuracy, accessibility, and potential sources of misinterpretation.
- Communicate insights from data, using visual representations to construct and support data-informed arguments that explain findings, acknowledge limitations, and address uncertainty.
- Identify and analyze sources of bias in datasets, models, and algorithms, and explain how these biases may affect conclusions or decisions.
- Evaluate ethical considerations in data practices (e.g., privacy, consent, transparency, and responsible data stewardship).
- Assess the societal impacts of data-driven systems, and discuss how data science can both reflect and shape social outcomes.
- Make reasoned judgments about responsible data use, considering, for example, fairness, accountability, and the limitations of data and algorithms.

Unit 4: Shuffling Songs: Probabilistic Modeling

Student Learning Outcomes:

- Describe and interpret variability and distributions in data using graphical displays and summary statistics, including measures of center and spread, consideration of outliers, and how distribution shape and context inform conclusions.
- Use probability and simulation to reason about randomness, estimate the likelihood of events, and interpret probabilistic outcomes in context.
- Analyze relationships between variables using visualizations and statistical measures to describe patterns, direction, and strength of association, including interpreting correlation and R2 to evaluate how well a model explains variability in context.
- Develop and evaluate statistical models to represent relationships in data, including selecting and interpreting appropriate function types (e.g., linear, exponential, or other

nonlinear models) and recognizing the usefulness and limitations of models in representing real-world phenomena.
- Use computational tools to process, analyze, and visualize data.
- Apply algorithms or basic programming to clean, transform, and explore datasets.
- Apply computational methods to support statistical analysis and modeling, and use digital tools to implement calculations, visualizations, and model fitting.
- Construct and document reproducible data analysis workflows by organizing computational steps and clearly communicating processes and decisions so analyses can be repeated, interpreted, and critiqued by others.

Unit 5: Skin Tones and Representation: Categorical Data and Introduction to Linear Algebra

Student Learning Outcomes:

- Evaluate the quality, credibility, and relevance of data sources for a given investigation.
- Analyze data collection processes, including sampling methods, representativeness, potential bias, reliability, and ethical considerations such as privacy, participant harm, and informed consent.
- Clean and prepare real-world datasets by addressing missing values, duplicates, inconsistent formats, and incomplete or unstructured data.
- Organize and transform data to support analysis, including joining datasets, restructuring variables, and converting data into usable formats.
- Describe and interpret variability and distributions in data using graphical displays and summary statistics, including measures of center and spread, consideration of outliers, and how distribution shape and context inform conclusions.
- Use probability and simulation to reason about randomness, estimate the likelihood of events, and interpret probabilistic outcomes in context.
- Analyze relationships between variables using visualizations and statistical measures to describe patterns, direction, and strength of association, including interpreting correlation and R2 to evaluate how well a model explains variability in context.
- Develop and evaluate statistical models to represent relationships in data, including selecting and interpreting appropriate function types (e.g., linear, exponential, or other nonlinear models) and recognizing the usefulness and limitations of models in representing real-world phenomena.

Unit 6: What's the Best Place for Me: Modeling with Data and Understanding Bias

Student Learning Outcomes:

- Evaluate the quality, credibility, and relevance of data sources for a given investigation.

- Analyze data collection processes, including sampling methods, representativeness, potential bias, reliability, and ethical considerations such as privacy, participant harm, and informed consent.
- Clean and prepare real-world datasets by addressing missing values, duplicates, inconsistent formats, and incomplete or unstructured data.
- Organize and transform data to support analysis, including joining datasets, restructuring variables, and converting data into usable formats.
- Identify and analyze sources of bias in datasets, models, and algorithms, and explain how these biases may affect conclusions or decisions.
- Evaluate ethical considerations in data practices (e.g., privacy, consent, transparency, and responsible data stewardship).
- Assess the societal impacts of data-driven systems, and discuss how data science can both reflect and shape social outcomes.
- Make reasoned judgments about responsible data use, considering, for example, fairness, accountability, and the limitations of data and algorithms.

Unit 7: Predicting My Preferences: Introduction to Machine Learning

Student Learning Outcomes:

- Use computational tools to process, analyze, and visualize data.
- Apply algorithms or basic programming to clean, transform, and explore datasets.
- Apply computational methods to support statistical analysis and modeling, and use digital tools to implement calculations, visualizations, and model fitting.
- Construct and document reproducible data analysis workflows by organizing computational steps and clearly communicating processes and decisions so analyses can be repeated, interpreted, and critiqued by others.
- Develop predictive models using machine learning tools, and use selected features and existing data to generate predictions about future or unknown outcomes.
- Evaluate predictive models using performance metrics (e.g., accuracy, error, or other indicators of model performance).
- Analyze model behavior and uncertainty, including recognizing situations of overfitting, underfitting, and limitations in predictions.
- Critically interpret predictions generated by models and implications of automated decision-making.

Unit 8: Being a Data Scientist

Student Learning Outcomes:

- Conduct data investigations as an iterative process, including: formulating statistical questions, exploring and analyzing data, revising assumptions or approaches, and communicating evidence-based conclusions.
- Use exploratory data analysis to identify patterns, test initial ideas, and determine next steps in a data investigation.

- Create and interpret purpose-focused data visualizations, and select appropriate representations for univariate, bivariate, and multivariate data.
- Evaluate and critique data visualizations for clarity, accuracy, accessibility, and potential sources of misinterpretation.
- Communicate insights from data, using visual representations to construct and support data-informed arguments that explain findings, acknowledge limitations, and address uncertainty.