

Baseline Professional Learning Handout

M212: Predictive Modeling

Baseline Professional Learning Handout	1
Selected Common Core High School Standards for Mathematics	2
Data Science Big Ideas	2
M212 Content and Practice Expectations	3
M212 Content and Practice Expectations and Indicators	4
Let's Investigate	6
Filtering Method Flowchart	6
Your Favorite Sandwich	7
Let's do some math! Round 1	8
Task 1: Visualizing Relationships	8
Learning Principles	10
Let's do some math! Round 2	11
Task 2: Decision Tree	11
Criteria for Success	13

Selected Common Core High School Standards for Mathematics

Calculate expected values and use them to solve problems

1. (+) Define a random variable for a quantity of interest by assigning a numerical value to each event in a sample space; graph the corresponding probability distribution using the same graphical displays as for data distributions.
2. (+) Calculate the expected value of a random variable; interpret it as the mean of the probability distribution.
3. (+) Develop a probability distribution for a random variable defined for a sample space in which theoretical probabilities can be calculated; find the expected value. For example, find the theoretical probability distribution for the number of correct answers obtained by guessing on all five questions of a multiple-choice test where each question has four choices, and find the expected grade under various grading schemes.
4. (+) Develop a probability distribution for a random variable defined for a sample space in which probabilities are assigned empirically; find the expected value. For example, find a current data distribution on the number of TV sets per household in the United States, and calculate the expected number of sets per household. How many TV sets would you expect to find in 100 randomly selected households?

Use probability to evaluate outcomes of decisions

5. (+) Weigh the possible outcomes of a decision by assigning probabilities to payoff values and finding expected values.
 - a. Find the expected payoff for a game of chance. For example, find the expected winnings from a state lottery ticket or a game at a fastfood restaurant.
 - b. Evaluate and compare strategies on the basis of expected values. For example, compare a high-deductible versus a low-deductible automobile insurance policy using various, but reasonable, chances of having a minor or a major accident.
6. (+) Use probabilities to make fair decisions (e.g., drawing by lots, using a random number generator).
7. (+) Analyze decisions and strategies using probability concepts (e.g., product testing, medical testing, pulling a hockey goalie at the end of a game).

Data Science Big Ideas

M212 Content and Practice Expectations

212.a	Engage in the modeling cycle.
212.b	Construct questions that can be answered by predictive modeling.
212.c	Select appropriate types of predictive models that help answer statistical questions and justify these choices.
212.d	Demonstrate understanding that $\text{Data} = \text{Model} + \text{Error}$, where the Model may contain one predictor, multiple predictors, or no predictors at all.
212.e	Understand and implement machine learning methods, and distinguish between supervised and unsupervised learning methods.
212.f	Improve predictive model fit by transforming variables, selecting variables, or using an alternate type of predictive modeling method.
212.g	Use predictive models to make predictions and interpret these predictions in context.
212.h	Distinguish between correlation and causation.

M212 Content and Practice Expectations and Indicators

Content and Practice Expectations	Indicators Choose an artifact where you...
212.a: Engage in the modeling cycle (performance assessment).	i. engage with the full modeling cycle (problem, formulate, compute, interpret, validate, revise as necessary, report).
212.b: Construct questions that can be answered by predictive modeling.	i. differentiate between statistical and non-statistical questions by recognizing that a statistical question can be answered by collecting data and anticipates variability in those data.
	ii. formulate statistical questions with the intent of answering the question using a predictive model.
	iii. postulate statistical questions that produce a given data visualization/model to answer those questions.
	iv. identify and define independent, dependent, and codependent relationships between potential variables in a given context or data set.
212.c: Select appropriate types of predictive models that help answer statistical questions and justify these choices.	i. identify the most appropriate model to use based on the objective of the statistical question, how the variables are measured, and type of data used (e.g., group versus regression models).
	ii. informally assess model fit visually and by comparing the coefficient of determination when applicable.
212.d: Demonstrate understanding that $\text{Data} = \text{Model} + \text{Error}$, where the Model may contain one predictor, multiple predictors, or no predictors at all.	i. define and use the concept of a statistical model as a function that produces a predicted score for each observation.
	ii. differentiate between the empty model and a more complex model.
	iii. demonstrate understanding of how to represent a linear relationship with a Least Squares Regression Line and interpret the slope and y-intercept in context.
	iv. understand that a perfect prediction does not usually exist by recognizing error visually.
212.e: Understand and implement machine learning methods, and distinguish between supervised and unsupervised learning methods.	i. build generalized linear models to fit data with one or more predictors, some of which may be transformed, and use those models to predict outcomes.
	ii. demonstrate understanding of which classification methods are used by implementing at least one type of classification method to make a prediction.

Content and Practice Expectations	Indicators Choose an artifact where you...
	iii. demonstrate understanding of which classification methods are used by implementing at least one type of clustering method to make a prediction.
212.f: Improve predictive model fit by transforming variables, selecting variables, or using an alternate type of predictive modeling method.	i. demonstrate understanding that some models visually appear to fit data better than other models, which can be confirmed by an increased coefficient of determination. ii. use dimensionality reduction techniques to decrease the number of variables in a predictive model. iii. produce a predictive model that has transformed variables.
212.g: Use predictive models to make predictions and interpret these predictions in context.	i. appropriately plug inputs into a predictive model to obtain an output, called a prediction. ii. explain what an output from a predictive model means in context, including the appropriate use of units when applicable. iii. identify limitations of a predictive model, including the trustworthiness of extrapolations and evaluating intercepts/slope in context, when applicable.
212.h: Distinguish between correlation and causation.	i. demonstrate understanding that correlated variables do not imply a causal relationship between those variables. ii. define confounding variables and describe the contextual meaning of an identified confounding variable.

Let's Investigate

Filtering Method Flowchart

Choose an app or website of your choice. As a group you will create an example flowchart of a filtering method for the app or website you chose. Use the questions below to develop your flowchart.

- What apps or websites do you use that recommend something to you?
 - Brainstorm a list of apps that have made a recommendation to you about what to listen to, read, watch, buy, etc.
- What does the app recommend? What form does the recommendation come in?
- How do you think the app makes the recommendation? Specifically, which filtering method do you think they use? How do you think the filtering method works?
 - Consider what the inputs might be and what the outputs (recommendations) are.
 - What do you think the app is doing with the inputs?
 - How do you think the inputs connect to the outputs?
 - What do you think are the different stages in the filtering process?

Create a flowchart of how the filtering method works on poster paper. Use arrows, labels, and color to make connections between the different stages of the filtering process, specifically, starting with the inputs and how they connect to the outputs (the recommendations).

Source: [YouCubed](#)

Your Favorite Sandwich

An algorithm is a step-by-step procedure for solving a problem, making something, or doing something. You could write an algorithm for just about anything like getting dressed for school, brushing your teeth, or making a sandwich.

- a. Write step-by-step instructions for making your favorite sandwich. What would the first five steps of the instructions be?
- b. Watch the [video](#) of a father testing his children's peanut butter and jelly sandwich making step-by-step instructions.
 - i. What do you notice and wonder about writing algorithms?
 - ii. How would you edit your step-by-step instructions?

Source: [YouCubed](#)

Let's do some math! Round 1

Task 1: Visualizing Relationships

Sample Learning Experience

In groups, ask students to create a visual representation that communicates the relationship between the following terms: correlated, spurious correlation, causation, random relationship, direct relationship, confounding variable, and mediating variable. Some examples [here](#).

As you walk around groups, encourage students to share their thinking with each other as they decipher between these terms and decide on a visual that they feel represents the situation. The process and conversation is more important than the final product.

After groups have created a visual, invite them to share their visual with the class. Highlight the many different ways that groups have represented the relationship. As a class, discuss the different representations and the meanings of the terms, considering how they relate to each other.

[Adapted from YouCubed.org](#)

Note: Lessons from the [YouCubed.org](#) website may require setting up a free account and logging in to access the whole lesson.

Learning Principles

Engage with cognitively demanding tasks in heterogeneous settings (LP 1). Students should be given opportunities to grapple with multistep, non-routine tasks that promote mathematical rigor. These experiences should be differentiated so that all students engage in appropriate challenges—for example, through tasks with multiple entry points and solution pathways. These experiences should continue to integrate knowledge and skills developed in grades 6-8 at the level of sophistication of high school mathematics.

Engage in social activities (LP 2). Students should have opportunities to work independently and to communicate with one another about mathematics by engaging in collective and collaborative learning activities. Explaining and having opportunities to revise one’s thinking has excellent value in solidifying one’s knowledge.

Build conceptual understanding through reasoning (LP 3). Students should be given the opportunity to reason, justify, and problem solve with critical thinking, reading, writing, speaking, and listening. Through reasoning and work with multiple representations, students learn why procedures work and build a conceptual understanding of key mathematical ideas.

Have agency in their learning (LP 4). Students should be able to choose tasks and learning experiences that align with their interests and aspirations. All students have rich and varied experiences and home lives. Learning mathematics should bring students’ identities and interests to the fore and build on the strengths that they bring to the learning space.

View mathematics as a human endeavor across centuries (LP 5). Students should understand that mathematical ideas emanated over time from civilizations around the world and have opportunities to explore these contributions to mathematics. Students should develop an appreciation of mathematics as a human endeavor: one in which they feel a sense of belonging, where they see themselves as mathematicians, and one that offers opportunities to broaden their ideas about what mathematics is, how it is used, and who it is for.

See mathematics as relevant (LP 6). Students should engage with mathematics in ways that authentically involve real-world situations. Problem-solving contexts should allow students to perceive mathematics as a tool for addressing the questions that arise in everyday life, as well as the ways it can model our world and address global economic, social, and environmental challenges. Students should also engage with mathematics in ways that connect to academic disciplines and future careers by doing the mathematics used by artists, designers, engineers, and other professionals.

Employ technology as a tool for problem-solving and understanding (LP 7). Research indicates that technology is a powerful tool for learning deeper mathematics by improving calculation efficiency and enabling more sophisticated analysis. Students should learn to use technology, with emphasis put on widely used tools and software, such as calculators and spreadsheets, to make sense of models. Technology use should not be limited to supporting “doing mathematics,” but should also be used as a tool for displaying and communicating results to appropriate audiences.

Task 1: Visualizing Relationships

Example 1:

Student work goes here

Example 2:

Student work goes here

Let's do some math! Round 2

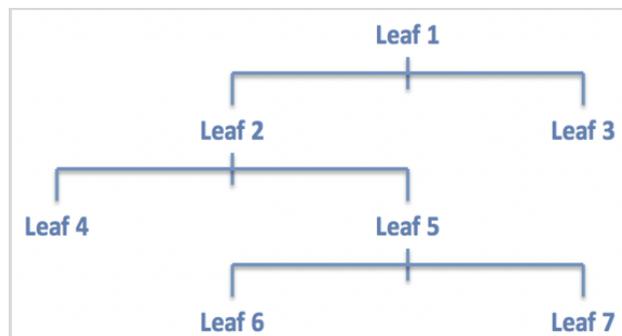
Task 2: Decision Tree

Students are given a prompt like this:

A decision tree is similar to a series of questions that are asked sequentially. Observations start by answering the first question (at the root of the tree), and then proceed along the different branches based on the answers they give to the questions that follow. At the end, based on all of the questions asked, observations are then classified as one of k classifications.

*Your goal today is to work with your group to design a decision tree that may help solve a classification problem. Your decision tree should have at least seven clearly defined **nodes/leaves** that either classify a person/object or prompt the need for a new leaf. Draw a diagram of your group's decision tree.*

Example decision tree structure:



Once your decision tree is complete, pair with a group to test out the effectiveness of each group's decision tree. Ask your partner group to think of a person/object that can be classified into one of your group's categories, and have them move down your decision tree until their person/object is classified. Identify if the person/object was properly classified, then take turns attempting to use your partner group's decision tree. Provide feedback on whether or not your person/object is properly classified.

Make updates to your decision tree based on your partner group's feedback and prepare to have a whole class discussion.

Sample Learning Experience

As students work on their decision trees, ask questions that help reinforce students' confidence in their choice of categories. Provide examples if students struggle to start (e.g., types of fruit, baseball player or basketball player, school subject, etc.).

Adapted from Introduction to [Data Science Curriculum](#)

After groups make final modifications to their decision trees, discuss as a class:

- *How do decision trees classify objects/people as being a member of a group?*
- *Did our decision trees make correct predictions every time?*
- *How can we figure out what questions to ask and in what order to minimize the number of 2 misclassifications?*

- *How is a decision tree similar to or different from a linear model?*
- *Can we call a decision tree a predictive model? Why or why not?*

Task 2: Decision Tree

Example 1:

Student work goes here

Example 2:

Student work goes here

Criteria for Success

Conference and Provide Revision Support	Accept with Revision	Accept
<p>The student's artifact shows an emerging understanding of the expectations of the indicator(s). After conferencing and additional instruction/learning, the student may provide a revised or different artifact as evidence of the indicator(s).</p>	<p>The student's artifact is approaching a full understanding of the expectations of the indicator(s). The artifact may contain execution errors that should be corrected in revision. The student may revise the selected artifact or submit a different artifact.</p>	<p>The student's artifact demonstrates evidence that they have met the expectations of the indicator(s).</p>