(http://csmatters.org/pd-new) C - 04a

0bC - 0b100

Data Acquisition: Modeling and Simulation (online)

Unit Concept Lessons

Revision Date: Sep 10, 2019
Duration: 120 50-minute sessions

Lesson Summary

Summary: Teachers will begin this lesson by journaling about how to create meaning out of data. They will then look at examples of nonsensical connections between data. After that, they will be introduced to data science. Teachers will think about the tradeoffs involved with data, as well as other uses of data. They will consider algorithms used with data collection and do a writing activity.

Outcomes:

Teachers will:

- Understand the basics of statistics and data science
- Understand the pros and cons of data, and how it is used in the modern world
- Learn the basics of algorithms as they relate to utilizing data in a meaningful way

Overview:

(Total: 120 minutes)

1. Journal activity
2. Investigating Data
    1. Gathering and Presenting Data
    2. Data Science as a Career
    3. Tradeoffs in gathering and analyzing data
    4. Algorithms that manipulate data
    5. Machine Learning and data
3. Discussion group post

Learning Objectives

Teacher Resources

CL04_Data Acquisition: Modeling and Simulation Folder (https://drive.google.com/open?id=0B5vAY-fhOT-idnFmNmdfb1d5d0U)

Lesson Plan

TOTAL: 120 minutes

Journal Activity

Describe at least 2 ways that we create meaning out of data.

- Possible answers include, but are not limited to: graph it, total, average, min and max, map it, find trends, generate predictions, etc.
- Do an online search and find out what results your students are most likely to come up with if you ask them to provide an answer to the question "How do we create meaning out of data?". Which answers make sense, and which don't?

Investigating Data

1. Gathering and Presenting Data

View the sample graph on slide 3 in the presentation on Data and Analysis. (https://docs.google.com/presentation/d/1GZOa8xlvYHZsQSPnvuJtPXatJNpqIcsrNF3aQHevnCY/edit#slide=id.g1b85463056_0_0)

- Subquestions:
  - Does eating margarine cause divorce in the state of Maine?
  - Does less divorce in Maine lead to fewer people buying margarine?
- What facts can we state from this data?

Search for some other examples of data connections that may be statistically valid, but the representations do not make sense.

2. Data science as a career

Watch the two videos explaining what a data scientist is and what a data scientist does on slide 4 in the presentation on Data and Analysis. (https://docs.google.com/presentation/d/1GZOa8xlvYHZsQSPnvuJtPXatJNpqIcsrNF3aQHevnCY/edit#slide=id.g1b85463056_0_0) (~3 min).

3. Tradeoffs in gathering and analyzing data

- EXAMPLE: Flying a space mission is a good example to use here. (One thing that has to be tested prior to space missions is the effects of sound on the equipment! Don't want anything breaking.)

4. Algorithms that manipulate data

- See if anyone knows how Netflix, movie makers, or Amazon use data about their customers to be more successful. Ask for a few opinions.
- See if anyone knows the story of Moneyball (based on a true story) of how a baseball team made decisions based on data analysis to become winners. (or watch the movie, it's good!)

Data analysis requires an algorithm. Consider what possible algorithm Netflix could use to suggest movies to a customer.

- Describe at least 2 calculations needed.
- Describe some of the data you would need to collect.
- What is a possible algorithm for making a decision about choosing: (pick one)
  - What movie to produce?
  - What sports player to hire?

5. Investigating algorithms

Machine learning in artificial intelligence is a very interesting use of data for making decisions. Watch as much as you can of the Stanford course on machine learning (https://www.youtube.com/watch?v=PPLop4L2eGk&list=PLLssT5z_DsK-h9vYZkQkYNWcItqhlRJLN) to get some background so you will be able to lead your students in a discussion about how machine learning works and the impacts of machine learning

Post to the discussion group:

1. How can we convince students that the real value of data depends on many factors: the accuracy, and timeliness of the data, the size and breadth of the sample, the appropriateness of the sample to the question, and the interpretation?
2. What aspects of data science might interest students since it is such a quickly growing area?
3. What considerations and tradeoffs arise in the computational manipulation of data?
4. Describe an algorithm that manipulates data
    1. Describe at least 2 calculations needed.
    2. Describe some of the data you would need to collect.
5. How do computers analyze data and learn from it?

 (http://www.umbc.edu/)  (http://www.umd.edu/)  (http://www.nsf.gov/)

Authored by: CS Matters in Maryland
Website: csmatters.org (http://csmatters.org)
Email: csmattersinmaryland@gmail.com (mailto:csmattersinmaryland@gmail.com)