

**Daily Lesson Plans****Instructional Days:** 1-3

**Topic Description:** This lesson sets the stage for the unit. It provides a review of how data can be used for making a case or as a vehicle for discovery. Students explore the pitfalls and challenges of putting together and managing large sets of data. An overview of the final project is provided.

**Objectives:**

The student will be able to:

- Explain the possible themes for the final project.
- Explain the difference between data used for making a case and data that informs discovery.
- Identify and discuss the considerations that must be made in order for a large data set to be useful.
- Consider how various types of data (numbers, text, dates, etc.) lend themselves to processing.
- Collaborate with others to create, manage, and maintain a large set of data.

**Outline of the Lesson:**

- Overview of final project. (35 minutes)
- Review of data collection and making a case/discovery. (5 minutes)
- Journal entry (15 minutes)
- Introduction to merging of diverse data sets about the same subject (10 minutes)
- Merge individual data sets together (45 minutes)
- Discussion of choices/decisions/compromises made during data merging (30 minutes)
- Discussion of creating an entire-class data set.(25 minutes)

**Student Activities:**

- Participate in discussion of final project.
- Participate in discussion of data review.
- Complete journal entry.
- Participate in discussion about data formats.
- Groups work to merge individual data sets together.
- Participate in discussion of choices/decisions/compromises made during data merging.
- Participate in discussion about creating an entire class data set.

**Teaching/Learning Strategies:**

- Introduce students to some possible research topics for the final project by using the suggested videos or other high-interest and engaging resources. The goal is to spark their interest and desire to learn more about the topic, and to brainstorm how they can be citizen scientists on the topics. The topics below are arranged from less to more complex.

- School safety and bullying
- Teen exercise and health
- Teens, social media, and online behavior
- Teens, video games, and civic engagement

Resources for the topics are included in Sample Final Project Topics. Additional resources related to these topics are included in later lessons.

- Journal Entry: In previous units we have discussed and considered many of the ways that data in the world are generated and collected. The term “Big Data” is often used to refer to the notion of very large data sets, and dealing with Big Data is something that many computer scientists concern themselves with. What does Big Data mean to you? In particular how big do you think a data set needs to be in order for it to qualify as Big Data? Do you think there is a formal definition? Why or why not?
  - Have students share their responses with their elbow partner.
  - Discuss student responses.
- Review the differences between data used for making a case and making a discovery.
- Data project—You can use the Room Data Project or create one of interest to your students that includes the following features and types of data:
  - Text
  - Spatial
  - Qualitative
  - Quantitative

and for which students can collect multiple entries by the next class period. You may wish to have them collect data related to one of the sample final project topics as an alternative. (The Room Data Project is adapted from a version by Baker Franke and Jeff Solin include in the *Taste of Computing* version of ECS.)

- Access to large data sets provides opportunities for in depth analysis, but in order to do this the data must be processed. In this project students will explore the challenges and pitfalls of putting together and managing a large set of data. Many choices can be made with no clear right or wrong answers. The goal is for the class to develop a single large set of data to use, formulate questions about the data, and answer them using an analysis tool. Each class’s will be different depending on the choices they make.
- Assign data collection for homework to be discussed the next class period.
- Break class into groups of 4 to merge their personal data sets into one.
  - Individuals in each group should share their school/room data spreadsheets with each other and discuss the choices they made.
    - How was Happiness recorded?
    - How were dates/times written?
    - How were locations recorded?
    - Did everyone record the same number of entries for each day?
    - If the goal is to merge the group’s data into one consistent set, how might the number of entries per student be affected?
  - It is likely that each person recorded their data differently.
  - The group’s task is to merge their four data sets into one consistent data set.

- The group should agree on how to denote days, times, locations and happiness.
  - It may be that some data needs to be compromised in order to conform to the group standard.
  - UNDER NO CIRCUMSTANCES should new data be “invented” after the fact to fit the mold.
- The group should also agree on how many records (rows) of data should be allocated for each person.
  - It need not be exactly the same number per person, but it should be representative of the group
  - For example if one student has a minute-by-minute account of their day, while another just made one entry for each class they were in, how can they make that consistent?
- Once the group agrees on a standard, they must produce one file that represents everyone’s data from the group
  - Students should use whatever collaboration technology is available to divide up the work of doing this - you shouldn’t have one student re-typing all the data. A “low-tech” suggestion is for each student to re-type/convert their personal data, save the file and send it to one person who can copy/paste the sheets together.
  - Note: this activity can also be done on poster paper without actual data entry, if desired.
- By the end of class each group should have saved a well-formatted file containing the group’s data.
- For homework students should now collect data in the group’s agreed upon format.
- Groups share their merged data (e.g. printing, projector, online tool)
  - Have a member from each group explain some of the choices that they made, explain challenges they overcame.
  - Guiding questions:
    - Is it important or not that the resulting file be consistent?
    - What does consistency in the data mean?
    - Why might it be important or why might it not be important?
    - What about differences in the number of entries per student? How might that affect the data?
- Have each group develop an interesting question (or two) to ask of the data
  - Examples:
    - What is the happiest room in school? least happy?
    - What are the top 3 visited rooms by people in this class?
    - What time of day is everyone busiest?
    - What day of the week are people happiest? least happy?
- Introduction to thinking about big data.
  - Have groups share their questions and answers—foreshadow the need for a larger set of data.
  - These are questions that we can ask when we have a large set of data to work with.
- As an entire class discuss how to merge all groups’ data into one data set.

- Goal: For each of the four types of data in the Room Data Collection Project, the entire class must agree on an appropriate format that will make processing the data to answer questions of it possible.
- For homework students should now collect data in the class' agreed upon format.

**Resources:**

- Room Data Collection Project

Sample Final Project Topics