



Lesson 17: Developing a Statistical Project

Student Outcomes

- Students construct a statistical question and a plan to collect data to answer the question.
- Given a statistical question, students use data to construct appropriate graphical and numerical summaries.
- Students use graphical and numerical summaries to answer a statistical question.

Lesson Notes

This lesson is an exploratory lesson. The following agenda provides an overview of the events of this lesson:

- Part 1: Review statistical questions posed in this module. (Approximately 10 minutes)
- Part 2: Summarize the four-step process that was used to begin this module. (Approximately 10 minutes)
- Part 3: Organize students in small groups to evaluate statistical questions. (Approximately 10 minutes)
- Part 4: Complete the four-step outline that is provided in the lesson. (Approximately 15 minutes)

In this lesson, students implement the four-step investigative process based on their own statistical question and data they collect. Students provide a plan to collect data to answer their statistical question. After their plans have been approved, students collect and organize their data and proceed to summarize the data using graphical and numerical summaries. The activity in this lesson is designed to provide a review of the four-step process that outlines for students a statistical study. This activity will require between 30 to 45 minutes of class time.

Students are expected to provide a daily update regarding their progress in collecting and summarizing the data they propose in this lesson. Students should have completed the numerical and graphical summaries before they start Lesson 21. Lesson 21 reviews the four steps and directs students to create a poster (or an outline for a presentation) based on their question, the data collected, and the numerical and graphical summaries. Lesson 22 outlines what is expected in their posters or presentations. Students conclude Lesson 22 by explaining their statistical project to their classmates or invited guests.

Several statistical questions from previous lessons launch this lesson. Direct students to read the list of statistical questions that they studied in this module. After they read this list and the two questions that follow, have a discussion based on the two summary questions.

Classwork

Statistical questions you investigated in this module included the following:

- How many hours of sleep do 6th graders typically get on a night when there is school the next day?
- What is the typical number of books read over the course of 6 months by a 6th grader?
- What is the typical heart rate of a student in a 6th grade class?
- How many hours does a 6th grader typically spend playing a sport or a game outdoors?
- What is the head circumference of adults interested in buying baseball hats?
- How long is the battery life of a certain brand of batteries?
- How many pets do students have?
- How long does it take a student to get to school?
- What is a typical daily temperature of New York City?
- What is the typical weight of a backpack for students at a certain school?
- What is the typical number of french fries in a large order from a fast food restaurant?
- What is the typical number of minutes a student spends on homework each day?
- What is the typical height of a vertical jump for a player in the NBA?

What do these questions have in common?

Why do several of these questions include the word “typical”?

Two discussion questions are posed to highlight the characteristics of a statistical question that was initially developed in Lesson 1 and continued throughout the lessons of this module.

- What is common about these questions?

If necessary, remind students that answering each of these questions require data. It is also anticipated that the data would vary, making them statistical questions.

- Why do several of these questions include the word “typical”?

Answering statistical questions often involves describing a data distribution. Often we are interested in finding a single value that describes what is typical of the values in a data set. Remind students that asking what value is typical for a group is a statistical question, whereas asking about the value of a single observation or about a single individual is not a statistical question because it is not a question that would be answered by collecting data that vary.

A Review of a Statistical Study

MP.4

To provide students an opportunity to see the entire four-step process, review a statistical study based on one or more of the questions they studied in the previous lessons and listed in the introduction of the lesson. A table to structure the discussion of the first three steps of the process is provided in the student's material. First, encourage students to select at least one of the statistical questions given and write it in the first row of the table. Next, students should recall the data collected to answer this question and to think about how the data might have been collected. Third, review the types of numerical summaries and graphs that students were either given or were expected to construct. Generally, students are expected to start with a graphical summary of the data. Then, based on the shape of the data distribution, the mean and MAD or the median and IQR would be used to describe center and spread. Finally, have students use the summaries to answer the statistical question. Again, reviewing one of the investigations from the lessons provides an opportunity to connect a conclusion to the other three steps.

The following table is a completed example that uses the question, "How many hours of sleep do 6th graders typically get on a night when there is school the next day?"

Recall from the very first lesson in this module that a statistical question is a question answered by data that you anticipate will vary.

Let's review the steps of a statistical investigation.

- Step 1: Pose a question that can be answered by data.
- Step 2: Collect appropriate data.
- Step 3: Summarize the data with graphs and numerical summaries.
- Step 4: Answer the question posed in Step 1 using the numerical summaries and graphs.

The first step is to pose a statistical question. Select one of the above questions and write it in the following Statistical Study Review Template.

The second step is to collect the data. In all of these investigations, you were given data. How do you think the data for the question you selected in Step 1 was collected? Write your answer in the summary below for Step 2.

The third step involves the various ways you summarize the data. List the various ways you summarized data for Step 3.

Step 1: Statistical question.

How many hours of sleep do 6th graders typically get on a night when there is school the next day? (Lesson 3)

Step 2: Collect data.

Students from a 6th grade class might have been asked to indicate how many hours they slept. The data would consist of the answers from all of the students.

Step 3: Summarize the data.

The first summary was to organize the data in a dot plot. This data set indicated a nearly symmetrical data distribution. Numerical summaries of the mean and the MAD would provide a description of the typical number of hours of sleep for 6th grade students and a measure of how much variability there was in the sleep times.

Finally, the fourth step is to answer the statistical question. The answer to the statistical question was the focus of the investigation in each of the lessons. Describing a data distribution in terms of shape, center, and spread or, depending on the shape of the data distribution, calculating the mean or the median of the data often answer statistical questions.

MP.3 The 4th and final step is the conclusion based on the summaries of Step 3. In the above example, students indicated that a typical number of hours of sleep for 6th graders were approximately 8.5 hours (the mean) and the mean absolute deviation was approximately one hour. Using the opening questions, complete as many of these summaries as necessary in order for students to understand what is expected of them in completing a statistical project.

Make sure students understand that the focus of each lesson was built around a question and that data were then summarized using graphs and the measures of center and variability. The process of collecting the data, however, was not a step students implemented in previous lessons. This lesson adds this new and important step into the investigative process. In general, collecting data will be done by asking students in your class to respond to the questions. In some cases, students may pose a question that could involve collecting data from other students in the school or from other friends or family connections. You might want to encourage some students to obtain data from an appropriate website. Monitor website choice by asking students to clearly indicate the site they plan to use. Check out the site to make sure students are getting appropriate data. A good site that was referenced in several lessons is the website of the American Statistical Association and their Census at School project (www.amstat.org/education/posterprojects/index.cfm).

This lesson should involve students collecting numerical data rather than categorical data. Indicate that students are expected to collect data that is numerical. Most of the questions that were used to start this lesson are possible questions that students could use to collect their own data. If necessary, review with students the difference between numerical and categorical data.

Often, students at this grade level do not appreciate data. They may construct a question that is unclear for those who are expected to answer it. For example, “How many hours of sleep do you get?” Several students answering that question would want to know what day of the week they should use. Help students clearly state their questions and avoid possibly collecting inaccurate data. Finally, make sure you screen the questions to make sure they are appropriate. If there are problems due to the immaturity of students in your class, remind them of the importance of getting their collecting plan approved. If problems continue, assist students in getting data from the Census at School project mentioned above or from other appropriate data sites.

Ideas for students to consider are provided in the opening questions of this lesson. Students could select one of these questions in their own statistical project. Review the winning posters from American Statistical Association’s poster competition (www.amstat.org/education/posterprojects/index.cfm) for additional ideas. Other ideas you may want to suggest during a discussion with students include the following:

- Number of languages spoken by teachers at your school.
- Height of students at your school (would require a sample of students as not every student attending the school could be included in the data collection).
- Number of words in the sentences of a Dr. Seuss book.
- Number of siblings for students in the school (would again require a sample of students).

Note: As indicated with some of the above suggestions, students may select a question that would require selecting a sample. Obtaining a good sample is an important part of statistical investigation. Students are introduced to random sampling and random samples in the seventh grade. For this project, however, challenge students to make sure that they consider ways in which a good representation of the population is collected. Although introducing students to a definition of random samples is not expected, students should be challenged to think about collecting a representative sample in order to answer their statistical question.

A formal problem set has not been added to this lesson. However, teachers are encouraged to design a problem set based on students’ progress during this lesson. The following options are possible ideas for setting up a problem set. Teacher discretion in organizing the project is important.

Option 1: Students who struggled with completing the three-step table developed around one of the questions launched with this lesson should be encouraged to select a different question and complete the table for this second question. The first three steps provide students a structure for connecting a question to a plan of collecting data and then summarizing the data. These steps were involved in the previous lessons; however, in this lesson, students need to bring the steps together. Once students organize these steps for a given question, they are ready to formulate their own question and data collection plan.

Option 2: Students provide a question and a data collection plan to teacher for review as outlined in the lesson. Direct students to complete the three-step table for their question and plan. Using the table provided in this lesson is an excellent way for students to organize their progress, plus it provides a good record for the teacher of how students are thinking at the beginning of this project. For students ready to begin this process, direct them to provide a summary of their statistical question, what plan they have for collecting the data, and what summaries of the data they anticipate will be done. Although these steps were discussed in the lesson, organizing this into a table similar to the one presented in the lesson provides you a summary of their progress. Periodically ask students during the next several days to update you on their progress by providing a summary of the table used in this lesson.

Option 3: For students going beyond the scope of questions outlined in the launch, they need to provide specific descriptions of what they plan to research and how they plan to collect the data. For example, if students explore research on honeybees, make sure students clearly indicate the statistical question (or what is the question that will be answered by data that is anticipated to vary), where they plan to obtain data to summarize the question, and how they plan to summarize the data. A brief written report or summary of their progress might constitute a workable problem set option for students at this level.

Project (Exploratory Challenge)

Now it is your turn to answer a statistical question based on data you collect. Before you collect data, explore possible statistical questions. For each question, indicate data that you would collect and summarize to answer the question. Also indicate how you plan to collect the data.

Think of questions that could be answered by data collected from members of your class or school or data that could be collected from recognized websites (e.g., The American Statistical Association and the project Census at Schools). Check with your teacher if you are planning to work with data from an outside source such as one of the above websites. Your teacher will need to approve both your question and your plan to collect data before data are collected.

As a class, explore possibilities of a statistical investigation. Record some of the ideas discussed by your class using the following table.

Possible statistical questions	What data would be collected and how would the data be collected?

After discussing several of the above possibilities of a statistical project, prepare a statistical question and a plan to collect data to present to your teacher. After your teacher approves your question and data collection plan, begin collecting the data. Carefully organize your data as you begin developing the summaries to answer your statistical question. In future lessons, you will be directed to begin creating a poster or an outline of a presentation that will be shared with your teacher and other members of your class.

For this lesson, complete the following to present to your teacher:

1. The statistical question for my investigation is as follows:
2. Here is the plan I propose to collect my data. (Include the exact questions you may ask an individual or a clear description of what you plan to measure or count.)

Lesson Summary

A statistical study involves a four-step investigative process:

- Pose questions that can be answered by data.
- Design a plan for collecting appropriate data and then use the plan to collect data.
- Analyze the data.
- Interpret results and draw valid conclusions from the data to the question posed.

Problem Set Sample Solutions

Your teacher will outline steps you are expected to complete in the next several days to develop this project. Keep in mind that the first step in developing your project is a statistical question. With one of the statistical questions posed in this lesson or with a new one developed in this lesson, organize your question and plan to collect and summarize data. Complete the process as outlined by your teacher.