

Click It

A-E Strand(s): Probability and Statistics. Sample Courses: Middle School Course 1, Middle School Course 2, Middle School One-Year Advanced Course, Integrated 3, and Algebra II.

Topic/Expectation

PS.D.1 Sample surveys, experiments and observational studies

- a. Describe the nature and purpose of sample surveys, experiments and observational studies, relating each to the types of research questions they are best suited to address.
- b. Recognize and explain the rationale for using randomness in research designs.
- d. Plan and conduct sample surveys to estimate population characteristics and experiments to compare treatments.

PS.3A Question formulation and data collection

- a. Formulate questions about a phenomenon of interest that can be answered with data.
- b. Design a plan to collect appropriate data.
- c. Collect and record data.

Core Algebra II EOC Content

EOC: S.2 Interpretation and communication through data

- c. Make judgments regarding accuracy, reasonableness, and bias in the use of data.
- d. Critique and justify various methods of sampling and data collection used in real world problems.

Rationale

Statistics is a new emphasis in most secondary mathematics programs. Among those programs that cover statistics, many tend to focus primarily on data analysis. This task goes beyond analyzing data by connecting the various discrete steps involved in designing and conducting a study. The task allows students to see the big picture of statistical studies and the various steps involved in undertaking such a study: identifying a question; determining how and from where or whom to gather data; conducting the study; gathering, organizing, and displaying the data; and analyzing and reporting results in response to the question.

Instructional Task

The National Highway Traffic Safety Administration (NHTSA) collects data about seat belt use and its relationship to age of driver, injury, economic costs, and so on. The NHTSA has reported some of its findings related to teenage drivers in a 2003 report called *The Facts To Buckle Up America: Safety Belts and Teens, 2003 Report*, available online at www.nhtsa.dot.gov/people/injury/airbags/buckleplan/buasbteens03/index.htm. The report includes the following paragraph:

Teens* have the highest fatality rate in motor vehicle crashes than any other age group.¹ There are many reasons; for instance, while teens are learning the new skills needed for driving, many frequently engage in high-risk behaviors, such as speeding and/or driving after using alcohol or drugs. Studies also have shown that teens may be easily distracted while

driving.² One key reason for high traffic fatalities among this age group is that they have lower safety belt use rates than adults.³ Because teens have an increased exposure to potentially fatal traffic crashes, it is imperative that efforts to increase safety belt use among this age group be given the highest priority. In addition, the youth population has increased by more than 12 percent since 1993, and is expected to increase by another seven percent by 2005.⁴ As this age group increases as a percentage of the population, the personal and societal costs associated with deaths and injuries from motor vehicle crashes also will rise.

References:

1. Traffic Safety Facts 2001 (Book), National Highway Traffic Safety Administration, DOT HS 809 484, pg. 21.
2. Williams, Alan F. 2001. Teenage Passengers in Motor Vehicle Crashes: A Summary of Current Research. Insurance Institute for Highway Safety.
3. NHTSA Research Notes, August 2001. DOT HS 809 318.
4. U.S. Census Bureau.

* For the purposes of this fact sheet, the term “teen” refers to young people ages 16–20 unless otherwise specified.

1. Discuss the statements made in the paragraph above. Which of the statements appear to report the findings of studies, and which appear to be opinions or analyses of the findings?
2. Which statements above describe or imply a causal relationship? For any such statements, describe the relationship and discuss how you might determine whether the implied or stated cause-effect relationship is an appropriate conclusion from the data.
3. The statement “the youth population has increased by more than 12 percent since 1993, and is expected to increase by another seven percent by 2005” is based on data from the U.S. Census Bureau. The government has strict guidelines about the statistical processes and sampling methods the Census Bureau can use. After exploring the Census Bureau’s website or investigating other reliable sources, describe the type of sampling the Census Bureau used in the 2000 census. How does this method allow the Census Bureau to deal with issues related to sampling and randomness? What advantages does its sampling procedure have over other types of sampling, such as systematic, stratified, or cluster sampling? What disadvantages might arise from the Census Bureau’s sampling process?
4. Construct a question related to teenage seatbelt use in your high school that a statistical study might help you answer. You might consider the factors listed above and/or other factors, such as grades, gender, smoking habits, or any other factors you think might be relevant.
5. Describe the experiment you will design to answer your question. Consider what population you would like to study or survey and discuss the sampling procedure you will use. Will you need to address randomization or other sampling issues, and if so, how will you do so? Justify your choice of sampling procedure.

Click It

6. Describe how you will gather data and how you will organize, display, report, and analyze these data for an audience you think would be interested in the question.
7. Conduct your study as described above and present your findings to an appropriate audience.

Discussion/Guiding Questions/Extensions

Teaching Notes

This task describes a long-term activity that could take several days or be assigned as an out-of-class project. The teacher might choose to have students work in groups. Students might be encouraged to use appropriate computer software or graphing calculator functions to present their data.

Extensions

A teacher might choose to have students conduct extended research about how the sampling procedures used for the US Census have changed over the years and what sampling procedures the Census Bureau will use for the 2010 census. The Census Bureau (census.gov) website has numerous statistical resources and provides free downloadable teaching materials. Within the site, the Survey Basics section for the American Community Survey (www.census.gov/acs/www/SBasics) provides a nice explanation of how this part of the census study is undertaken. Other discussions of the census sampling process and the census survey plan (such as “Census Sampling Confusion” in *Science News Online*, available at www.sciencenews.org/pages/sn_arc99/3_6_99/bob1.htm).

The study of sampling in the context of the census is an opportunity for cross-disciplinary work related to government as well as to discussions about democracy and fairness. The Supreme Court has made decisions affecting the reliability of census data, and the economic, political and social consequences of census data collected are significant for individuals, communities, and the nation.

Sample Solutions

1. Discuss the statements made in the paragraph above. Which of the statements appear to report the findings of studies, and which appear to be opinions or analyses of the findings?

Answers may vary. The paragraph seems to be a summary of various studies and reports. The first statement (“Teens...”) reports direct data. The next statement (“There are...”) appears to present an opinion, and the next (“Studies also...”) seems to report results. The statement beginning with “One key reason...” appears to be an attribution of cause and effect, or an analysis of some finding, while the next statement (“Because teens...”) is a recommendation based on an analysis of the various studies summarized. The statement beginning with “In addition...” appears to be direct report of census data, while the concluding sentence (“As this age group...”) is an inference or conclusion extrapolated from the data.

2. Which statements above describe or imply a causal relationship? For any such statements, describe the relationship and discuss how you might determine whether the implied or stated cause-effect relationship is an appropriate conclusion from the data.

As discussed in the answer to question 1 (above), the statement beginning with “One key reason...” appears to be an attribution of cause and effect, or an analysis of some findings.

3. The statement “the youth population has increased by more than 12 percent since 1993, and is expected to increase by another seven percent by 2005” is based on data from the U.S. Census Bureau. The government has strict guidelines about the statistical processes and sampling methods the Census Bureau can use. After exploring the Census Bureau’s website or investigating other reliable sources, describe the type of sampling the Census Bureau used in the 2000 census. How does this method allow the Census Bureau to deal with issues related to sampling and randomness? What advantages does its sampling procedure have over other types of sampling, such as systematic, stratified, or cluster sampling? What disadvantages might arise from the Census Bureau’s sampling process?

The U.S. Census was originally designed to be a count or interview of every household. Thus, it could be considered to have eliminated the need for a sample since it surveyed the entire population being studied. However, such a data-collection method introduces practical issues of whether every person can actually be counted, as well as broader issues related to one person reporting for others in a household, the accuracy of self-reported data, and the return rate of surveys. Consequently, some changes were made in how the data were collected for the 1990 census and again for the 2000 census to allow some sampling to be used as a supplement to the data gathered directly from households. Sampling issues are complex, however, and a different structure is planned for 2010. Sampling issues are discussed on the websites provided under the Extensions section.

Questions 4 through 7 should be assessed individually for each student study. The American Statistical Association’s (ASA) *Guidelines for Assessment and Instruction in Statistics Education*, ASA, 2006 (www.amstat.org/education/gaise/) can serve as a resource for evaluating students’ studies.