



AP[®] Statistics

Course Overview

About the Advanced Placement Program[®] (AP[®])

The Advanced Placement Program[®] has enabled millions of students to take college-level courses and earn college credit, advanced placement, or both, while still in high school. AP Exams are given each year in May. Students who earn a qualifying score on an AP Exam are typically eligible, in college, to receive credit, placement into advanced courses, or both. Every aspect of AP course and exam development is the result of collaboration between AP teachers and college faculty. They work together to develop AP courses and exams, set scoring standards, and score the exams. College faculty review every AP teacher's course syllabus. We commit to supporting educators and communities in their efforts to make AP courses widely available, advancing students in their plans for college and careers.

AP Statistics Course Overview

The AP Statistics course introduces students to the major concepts and tools for formulating questions, collecting and analyzing data, and interpreting results from data. The practices and skills for the course have been aligned to help students understand the statistical problem-solving process based on the American Statistical Association (ASA) recommendations. The content, skills, and assessments in the AP Statistics course focus on exploring data, sampling and experimentation, probability and simulation, and statistical inference. Students use technology, investigations, problem solving, and writing as they build conceptual understanding.

The AP Statistics course is equivalent to a one-semester, introductory, non-calculus-based college course in statistics.

The AP Statistics course is an excellent option for any secondary school student who has successfully completed a first-year algebra course.

AP Statistics Course Content

The course content is organized into units of study that provide a suggested sequence for the course. These units comprise the content and conceptual understandings in which colleges and universities typically expect students to be proficient to qualify for college credit and/or placement.

- **Unit 1:** Exploring One-Variable Data and Collecting Data
- **Unit 2:** Probability, Random Variables, and Probability Distributions
- **Unit 3:** Inference for Categorical Data—Proportions
- **Unit 4:** Inference for Quantitative Data—Means
- **Unit 5:** Regression Analysis

Each unit is broken down into teachable segments called topics.

Statistical Practices

- **Formulate Questions:** Determine an investigative question for a statistical study.
- **Collect Data:** Identify and justify methods for collecting data and conducting statistical inference.

- **Analyze Data:** Construct representations of data and calculate numerical statistical outputs.
- **Interpret Results:** Interpret results and justify conclusions and methods.

AP Statistics Exam Structure

AP STATISTICS EXAM: 3 HOURS

Assessment Overview

The AP Statistics Exam assesses student understanding of the statistical practices and learning objectives outlined in the course framework. The exam is 3 hours long and includes 42 multiple-choice questions and four 10-point free-response questions, each weighted equally and scored on an analytic scale. A graphing calculator with statistical capabilities is expected for both sections of the exam. Formulas and tables are provided for both sections of the exam.

Format of Assessment

Section I: Multiple Choice | 42 Questions | 90 Minutes |
50% of Exam Score

- Questions are comprised of discrete and set-based questions and may focus on scenarios related to a variety of work applications and/or real-life contexts.
- All five units of the course are assessed.
- All four statistical practices are assessed.

Section II: Free Response | 4 Questions | 90 Minutes |
50% of Exam Score

- **Question 1:** Multi-Focus on Practices 1 and 2
- **Question 2:** Multi-Focus on Practices 3 and 4
- **Question 3:** Inference (Hypothesis Test or Confidence Interval)
- **Question 4:** Multi-Focus on Practices 2, 3, and 4

Exam Components

Sample Multiple-Choice Questions

1. A one-sample z-test for a proportion at the significance level of $\alpha = 0.10$ will be conducted for the following set of hypotheses.

$$H_0: p = 0.50$$

$$H_a: p < 0.50$$

Which of the following statements represents a Type II error for the hypotheses?

- (A) Rejecting the null hypothesis when the population proportion is 0.50
- (B) Rejecting the null hypothesis when the population proportion is less than 0.50
- (C) Failing to reject the null hypothesis when the population proportion is 0.50
- (D) Failing to reject the null hypothesis when the population proportion is less than 0.50

Questions 2–4 refer to the following.

Emmalyn runs track-and-field at a local running club. To complete a project in her statistics class, she collects data from her teammates, measuring their 60-meter-dash time, in seconds, and long-jump distance, in inches. Using these data, she created a linear regression model using dash time as the x-variable and long-jump distance as the y-variable. The model she created is $\hat{y} = 313.875 - 30.805x$ and has $r = -0.591$ and $r^2 = 0.349$.

2. Which of the following is a correct interpretation of the correlation coefficient?
- (A) There is no linear relationship between dash time and long-jump distance.
 - (B) The linear relationship between dash time and long-jump distance is moderate and negative.
 - (C) There is a perfect linear relationship between dash time and long-jump distance.
 - (D) The linear relationship between dash time and long-jump distance is moderate and positive.
3. Which of the following is a correct interpretation for the coefficient of determination?
- (A) Approximately 35% of the variation in dash time can be explained by long-jump distance.
 - (B) Approximately 59% of the variation in dash time can be explained by long-jump distance.
 - (C) Approximately 35% of the variation in long-jump distance can be explained by the linear relationship with dash time.
 - (D) Approximately 59% of the variation in long-jump distance can be explained by the linear relationship with dash time.
4. A teammate with a dash time of 7 seconds had a residual of -6.8 inches. Based on the model, what is the observed long-jump distance of this teammate, in inches?
- (A) 91.44
 - (B) 98.24
 - (C) 105.04
 - (D) 313.88

Sample Free-Response Question

A country's transportation department wants to conduct a survey to collect data from adults, ages 18 and older, in that country. The survey will be focused on commuting to work, including commute time in minutes, method of transportation used, and other information.

Country A is divided into many regions, and each region will be asked to send out a survey and collect the data. The transportation department will allow each region to use any method it prefers to survey the adults.

Three different regions will use the following sampling methods:

Region X: Each adult who lives in the region will receive a survey and will be required to complete the survey by mail. Adults who do not complete the survey will be sent reminders to do so until they have completed the survey.

Region Y: The region is divided into six areas that are similar to each other. Two of the six areas will be randomly selected, and each adult who lives in those two areas will receive a survey and be asked to complete the survey by mail. The table gives the results of the data collection for the two areas in Region Y for the question "What is your average daily commute time to work in minutes?"

Region Z: Each adult who lives in the region and has responded to a previous official survey will receive a survey and be asked to complete it by mail.

Summary Statistics

Area	<i>n</i>	$\bar{x}$	Min	Q1	Median (Q2)	Q3	Max
One	2,887	55	7	35	53	74	101
Five	4,502	42	5	24	32	68	83

- A.
 - i. **Identify** the sampling method for Region X.
 - ii. **Identify** the sampling method for Region Y.
- B.
 - i. **Explain** why the results of the sampling method for Region Y can be generalized to the population of the region.
 - ii. Tom, a member of the planning committee for Region Z, proposed Region Z's sampling method. Tom claims the sampling method for Region Z may lead to a higher survey response rate than the sampling methods used for the other two regions and could better represent the entire population in Region Z. **Explain** why Tom's claim is not correct.
- C. Manuel, an insurance agent in Region Y, wants to compare the average commute times to work for the populations in each of the two areas of Region Y. His first step is to construct a graphical representation for each area using the summary statistics in the table.
 - i. **Identify** a graphical representation that is appropriate to use to display the summary statistics.
 - ii. **Explain** why the graphical representation from part C (i) is appropriate to use.
- D. Use comparisons of the summary statistics in the table to **describe** the most likely shape for the distribution of commute time in Area Five.
- E. Manuel wants to perform a study about the average daily commute time to work for Area Three and Area Six in Region Y. He believes there is a difference in the commute times for the two areas. **Determine** a valid investigative question for Manuel's study. Ensure all three components of a valid investigative question are included in your response.