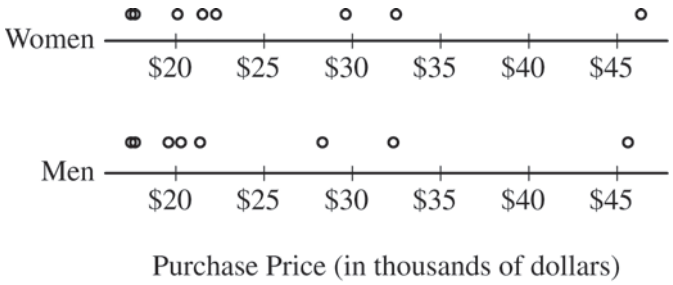
5. A researcher conducted a study to investigate whether local car dealers tend to charge women more than men for the same car model. Using information from the county tax collector's records, the researcher randomly selected one man and one woman from among everyone who had purchased the same model of an identically equipped car from the same dealer. The process was repeated for a total of 8 randomly selected car models.

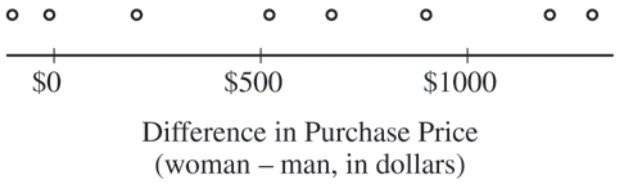
   The purchase prices and the differences (woman – man) are shown in the table below. Summary statistics are also shown.

Car model	1	2	3	4	5	6	7	8
Women	$20,100	$17,400	$22,300	$32,500	$17,710	$21,500	$29,600	$46,300
Men	$19,580	$17,500	$21,400	$32,300	$17,720	$20,300	$28,300	$45,630
Difference	$520	−$100	$900	$200	−$10	$1,200	$1,300	$670

	Mean	Standard Deviation
Women	$25,926.25	$9,846.61
Men	$25,341.25	$9,728.60
Difference	$585.00	$530.71

Dotplots of the data and the differences are shown below.





Do the data provide convincing evidence that, on average, women pay more than men in the county for the same car model?

© 2014 The College Board.
Visit the College Board on the Web: www.collegeboard.org.

STATISTICS

SECTION II

Part B

Question 6

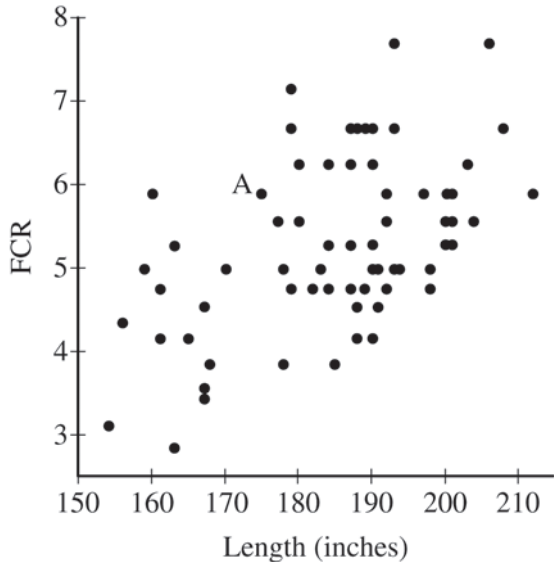
Spend about 25 minutes on this part of the exam.

Percent of Section II score—25

Directions: Show all your work. Indicate clearly the methods you use, because you will be scored on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

6. Jamal is researching the characteristics of a car that might be useful in predicting the fuel consumption rate (FCR); that is, the number of gallons of gasoline that the car requires to travel 100 miles under conditions of typical city driving. The length of a car is one explanatory variable that can be used to predict FCR. Graph I is a scatterplot showing the lengths of 66 cars plotted with the corresponding FCR. One point on the graph is labeled A.

GRAPH I



Jamal examined the scatterplot and determined that a linear model would be a reasonable way to express the relationship between FCR and length. A computer output from a linear regression is shown below.

Linear Fit
FCR = −1.595789 + 0.0372614 * Length

Summary of Fit	
RSquare	0.250401
Root Mean Square Error	0.902382
Observations	66

(a) The point on the graph labeled A represents one car of length 175 inches and an FCR of 5.88. Calculate and interpret the residual for the car relative to the least squares regression line.

© 2014 The College Board.
Visit the College Board on the Web: www.collegeboard.org.

AP® STATISTICS
2014 SCORING GUIDELINES

Question 5

Intent of Question

The primary goal of this question was to assess students' ability to identify, set up, perform, and interpret the results of an appropriate hypothesis test to address a particular question. More specific goals were to assess students' ability to (1) state appropriate hypotheses; (2) identify the appropriate statistical test procedure and check appropriate conditions for inference; (3) calculate the appropriate test statistic and *p*-value; and (4) draw an appropriate conclusion, with justification, in the context of the study.

Solution

Step 1: States a correct pair of hypotheses.

Let μ_{diff} represent the population mean difference in purchase price (woman – man) for identically equipped cars of the same model, sold to both men and women by the same dealer, in the county.

The hypotheses to be tested are $H_0 : \mu_{\text{diff}} = 0$ versus $H_a : \mu_{\text{diff}} > 0$.

Step 2: Identifies a correct test procedure (by name or by formula) and checks appropriate conditions.

The appropriate procedure is a paired *t*-test.

The conditions for the paired *t*-test are:

1. The sample is randomly selected from the population.
2. The population of price differences (woman – man) is normally distributed, or the sample size is large.

The first condition is met because the car models and the individuals were randomly selected. The sample size ($n = 8$) is not large, so we need to investigate whether it is reasonable to assume that the population of price differences is normally distributed. The dotplot of sample price differences reveals a fairly symmetric distribution, so we will consider the second condition to be met.

Step 3: Correct mechanics, including the value of the test statistic and *p*-value (or rejection region).

The test statistic is $t = \dfrac{585 - 0}{\dfrac{530.71}{\sqrt{8}}} \approx 3.12$.

The *p*-value, based on a *t*-distribution with $8 - 1 = 7$ degrees of freedom, is 0.008.

Step 4: States a correct conclusion in the context of the study, using the result of the statistical test.

Because the *p*-value is very small (for instance, smaller than $\alpha = 0.05$), we reject the null hypothesis. The data provide convincing evidence that, on average, women pay more than men in the county for the same car model.

© 2014 The College Board.
Visit the College Board on the Web: www.collegeboard.org.

AP® STATISTICS
2014 SCORING GUIDELINES

Question 5 (continued)

Scoring

Each of steps 1, 2, 3, and 4 were scored as essentially correct (E), partially correct (P), or incorrect (I).

Step 1 is scored as follows:

Essentially correct (E) if the response identifies the correct parameter *AND* states correct hypotheses.

Partially correct (P) if the response identifies the correct parameter *OR* states correct hypotheses, but not both.

Incorrect (I) if the response does not meet the criteria for E or P.

Note: Defining the parameter symbol in context or simply using common parameter notation is sufficient.

Step 2 is scored as follows:

Essentially correct (E) if the response identifies the correct test procedure (by name or by formula) *AND* checks *both* conditions correctly.

Partially correct (P) if the response correctly completes two of the three components (identification of procedure, check of randomness condition, check of normality condition).

Incorrect (I) if the response does not meet the criteria for E or P.

Note: The random sampling condition can be verified by referring to the random selection of car models or to the random selection of male and female car buyers.

Step 3 is scored as follows:

Essentially correct (E) if the response correctly calculates both the test statistic and the *p*-value.

Partially correct (P) if the response correctly calculates the test statistic but not the *p*-value;
OR
if the response calculates the test statistic incorrectly but then calculates the correct *p*-value for the computed test statistic.

Incorrect (I) if the response does not meet the criteria for E or P.

Note: If the response identifies a *z*-test for a mean as the correct procedure in step 2, then the response can earn a P in step 3 if both the test statistic and the *p*-value are calculated correctly.

© 2014 The College Board.
Visit the College Board on the Web: www.collegeboard.org.

Question 5 (continued)

Step 4 is scored as follows:

Essentially correct (E) if the response provides a correct conclusion in context, also providing justification based on linkage between the *p*-value and the conclusion.

Partially correct (P) if the response provides a correct conclusion with linkage to the *p*-value, but not in context;
OR
if the response provides a correct conclusion in context, but without justification based on linkage to the *p*-value.

Incorrect (I) if the response does not meet the criteria for E or P.

Notes:

- If the conclusion is consistent with an incorrect *p*-value from step 3 and also in context with justification based on linkage to the *p*-value, step 4 is scored as E.
- A response that performs a two-sample *t*-test with correct calculations should fail to reject H_0. A conclusion that is equivalent to "accept H_0" (such as "we conclude that women pay the same amount as men, on average"), either as a stated decision or as a conclusion in context, cannot be scored as E. Such a response will be scored as P provided that the conclusion is in context with linkage. Such a response will be scored as I if it lacks either context or linkage.

Each essentially correct (E) step counts as 1 point. Each partially correct (P) step counts as ½ point.

4 **Complete Response**

3 **Substantial Response**

2 **Developing Response**

1 **Minimal Response**

If a response is between two scores (for example, 2½ points), use a holistic approach to decide whether to score up or down, depending on the overall strength of the response and communication.

© 2014 The College Board.
Visit the College Board on the Web: www.collegeboard.org.