Answer Key for AP Statistics
Practice Exam, Section I

Question 1: C
Question 2: E
Question 3: C
Question 4: A
Question 5: A
Question 6: C
Question 7: C
Question 8: E
Question 9: B
Question 10: C
Question 11: B
Question 12: D
Question 13: B
Question 14: D
Question 15: B
Question 16: D
Question 17: D
Question 18: C
Question 19: C
Question 20: E
Question 21: C
Question 22: B
Question 23: C
Question 24: B
Question 25: A
Question 26: C
Question 27: B
Question 28: D
Question 29: B
Question 30: B
Question 31: C
Question 32: E
Question 33: B
Question 34: E
Question 35: E
Question 36: E
Question 37: D
Question 38: A
Question 39: C
Question 40: E

Multiple-Choice Section for Statistics
2019 Course Framework Alignment and Rationales

Question 1

Skill	Learning Objective	Topic
2.A	UNC-1.H	Describing the Distribution of a Quantitative Variable

(A)	Incorrect. The distribution is not approximately normal, since the distribution is neither mound shaped nor symmetric.
(B)	Incorrect. It is true that the distribution is bimodal. However, there are no observed data values between 1 and 8, so there is a gap displayed in the distribution.
(C)	**Correct**. The distribution is bimodal, with one mode at 10 and another mode at 17. Also, there are no observed data values between 1 and 8, so there is a gap displayed in the distribution.
(D)	Incorrect. The distribution is not skewed to the right. A distribution is skewed to the right when the right tail is longer than the left. However, there are no observed data values between 1 and 8, so there is a gap displayed in the distribution.
(E)	Incorrect. It is correct that there is a gap in the distribution. However, the distribution is not skewed to the right. A distribution is skewed to the right when the right tail is longer than the left tail.

Question 2

Skill	Learning Objective	Topic
3.A	VAR-4.D	Conditional Probability

(A)	Incorrect. This is the probability that the person selected is age 55 or older and responded no; it is not the probability that the person selected will be someone who responded no, given that the person selected is age 55 or older.
(B)	Incorrect. This is the probability that the person selected is age 55 or older; it is not the probability that the person selected will be someone who responded no, given that the person selected is age 55 or older.
(C)	Incorrect. This is the probability that the person selected was age 55 or older given that the person selected is someone who responded no; it is not the probability that the person selected will be someone who responded no, given that the person selected is age 55 or older.
(D)	Incorrect. This is the probability that the person selected answered no; it is not the probability that the person selected will be someone who responded no, given that the person selected is age 55 or older.
(E)	**Correct**. The condition given specifies that the person selected is age 55 or older, and this condition restricts the sample space to 44 people. Of those 44 people, 36 responded no, so the probability is found by $\frac{36}{44} \approx 0.818$.

Question 3

Skill	Learning Objective	Topic
2.A	DAT-1.F	Residuals

(A)	Incorrect. Point A in Graph 2 has a predicted fleece weight of approximately 7, not a predicted fleece weight of approximately 10, so it cannot be the residual for the circled point in Graph 1.
(B)	Incorrect. Point B in Graph 2 has a predicted fleece weight of approximately 8, not a predicted fleece weight of approximately 10, so it cannot be the residual for the circled point in Graph 1.
(C)	**Correct**. The circled point in Graph 1 corresponds to the sample value that has a fiber diameter of approximately 26 and a predicted fleece weight of approximately 10. For that point, the value of the residual fleece weight can be found using values for the observed fleece weight and predicted fleece weight from Graph 1. The value of the residual is given by $\text{residual} = \text{observed} - \text{predicted} \approx 5 - 10 \approx -5$. Point C is the point on Graph 2 that has a predicted fleece weight of approximately 10 and that has a residual fleece weight that is approximately -5.
(D)	Incorrect. Point D in Graph 2 has a predicted fleece weight of approximately 10, but a residual value of approximately -3, not -5, so it cannot be the residual for the circled point in Graph 1.
(E)	Incorrect. Point E in Graph 2 has a predicted fleece weight of approximately 10, but a residual value of approximately 5, not -5, so it cannot be the residual for the circled point in Graph 1.

Question 4

Skill		Learning Objective	Topic
2.A		UNC-1.H	Describing the Distribution of a Quantitative Variable
(A)	**Correct**. The only shape listed that is not represented by one of the distributions is a uniform shape. The shape of the weight distribution is bimodal. The shape of the pH distribution is skewed to the right. The shape of the flexibility rating distribution is skewed to the left. The shape of the octane rating distribution is symmetric and unimodal.		
(B)	Incorrect. The shape of the weight distribution is bimodal.		
(C)	Incorrect. The shape of the flexibility rating distribution is skewed to the left.		
(D)	Incorrect. The shape of the pH distribution is skewed to the right.		
(E)	Incorrect. The shape of the octane rating distribution is symmetric and unimodal.		

Question 5

Skill		Learning Objective	Topic
2.C		UNC-1.Q	Statistics for Two Categorical Variables
(A)	**Correct**. Of the 1,092 people who responded, 192 responded no to color consideration and also identified safety as the additional feature that is important. The proportion of people who responded no to color consideration and who identified safety as the additional feature that was important is $\frac{192}{1{,}092} \approx 0.18$.		
(B)	Incorrect. This is the proportion of the 1,092 people who responded that safety was the additional feature that was important.		
(C)	Incorrect. This is the proportion of the 534 people who responded no to color consideration who also identified safety as the additional feature that was important.		
(D)	Incorrect. This is the proportion of the 1,092 people who responded no to color consideration.		
(E)	Incorrect. This is the proportion of the 1,092 people who did <u>not</u> respond no to color consideration.		

Question 6

Skill	Learning Objective	Topic
3.A	VAR-2.B	The Normal Distribution

(A)	Incorrect. This is an age that is close to the age of a tortoise at the 10th percentile, not the 90th percentile, of the distribution.
(B)	Incorrect. This is an age that is close to the age of a tortoise at the 85th percentile, not the 90th percentile, of the distribution.
(C)	**Correct**. The value of approximately 119.22, found using technology, is the value that has 90 percent of the area to the left of it in the normal distribution with mean 100 and standard deviation 15. Of the values listed, 120 is the tortoise age that is closest to 119.22.
(D)	Incorrect. This is an age that is close to the age of a tortoise at the 95th percentile, not the 90th percentile, of the distribution.
(E)	Incorrect. This is an age that is close to the age of a tortoise at the 98th percentile, not the 90th percentile, of the distribution.

Question 7

Skill	Learning Objective	Topic
2.D	UNC-1.N	Comparing Distributions of a Quantitative Variable

(A)	Incorrect. Boxplots provide information on the proportion of values between certain measures in a distribution, but they give no information about the number of rentals for the locations.
(B)	Incorrect. Boxplots provide information on the proportion of values between certain measures in a distribution, but they give no information about the number of rentals for the locations.
(C)	**Correct**. There is more variability in the miles driven for location B than for location A since the interquartile range is greater for B than for A $(120 > 50)$ and the range of values for B is greater than the range of values for A. Also, the median number of miles driven is greater for location B than for location A $(80 > 50)$.
(D)	Incorrect. It is true that the median is greater for B than for A. However, the miles driven for location B display more variability, not less variability.
(E)	Incorrect. The miles driven for location B display more variability, not less variability, and the median is <u>not</u> about the same for B as it is for A.

Question 8

Skill		Learning Objective	Topic
1.C		DAT-2.C	Random Sampling and Data Collection
(A)	Incorrect. No experiment was conducted; the items and prices were observed and recorded.		
(B)	Incorrect. No experiment was conducted; the items and prices were observed and recorded.		
(C)	Incorrect. The end-of-year activity was not a sample survey, since no sample was selected; every item in stock was used.		
(D)	Incorrect. The end-of-year activity was not a sample survey, since no sample was selected; every item in stock was used.		
(E)	**Correct**. The end-of-year activity described is a census, since a list is made of every item in stock along with its corresponding wholesale price.		

Question 9

Skill	**Learning Objective**	**Topic**
3.A	VAR-2.B	The Normal Distribution

(A)	Incorrect. The z-score for the Ohio weight should be positive, so the number of standard deviations should be above the mean, not below the mean.
(B)	**Correct**. The number of standard deviations from the mean is given by $z = \frac{x - \mu}{\sigma}$. For the farm in Iowa, the z-score is 1.645, the value of x is 1.39, the value of μ is 1.26, and the value of σ is unknown. Thus, $1.645 = \frac{1.39 - 1.26}{\sigma}$, and solving for σ yields approximately 0.079. For the farm in Ohio, the value of σ is 0.01 greater than the value of σ for Iowa, so $\sigma = 0.079 + 0.01 = 0.089$. The z-score for Ohio is equal to $z = \frac{1.39 - 1.26}{0.089} \approx 1.46$, so the weight with respect to the Ohio distribution is 1.46 standard deviations above the mean.
(C)	Incorrect. The z-score for the Ohio weight was incorrectly calculated by using a standard deviation of 0.079; 0.089 should have been used.
(D)	Incorrect. The z-score for the Ohio weight was incorrectly calculated by using a standard deviation of 0.069; 0.089 should have been used. Also, the number of standard deviations should be above the mean, not below the mean.
(E)	Incorrect. The z-score for the Ohio weight was incorrectly calculated by using a standard deviation of 0.069; 0.089 should have been used.

Question 10

Skill	Learning Objective	Topic
3.A	VAR-6.B	The Normal Distribution, Revisited

(A)	Incorrect. This is the probability that the number of hours worked by a volunteer selected at random is greater than 90 in a normal distribution with mean 80 and standard deviation $7,$ not the probability that the volunteer selected will receive the certificate of merit given that the number of hours the volunteer worked is less than $90.$
(B)	Incorrect. This is the probability that a volunteer selected at random will have worked between 85.89 hours and 90 hours, not the probability that the volunteer selected will receive the certificate of merit given that the number of hours the volunteer worked is less than $90.$
(C)	**Correct**. If X represents the number of hours worked, then the value of X for which 20 percent of the hours worked are greater than X in a normal distribution with mean 80 and standard deviation 7 can be found using technology to be approximately $85.89.$ Then the probability that the volunteer selected will receive a certificate of merit given that the number of hours the volunteer worked is less than 90 is given by $P(X > 85.89 \mid X < 90) = \dfrac{P(85.89 < X < 90)}{P(X < 90)}.$ Technology can be used to find that $P(85.89 < X < 90) \approx 0.1235$ and that $P(X < 90) \approx 0.924$ in a normal distribution with mean 80 and standard deviation $7,$ so $P(X > 85.89 \mid X < 90) = \dfrac{P(85.89 < X < 90)}{P(X < 90)} \approx \dfrac{0.1235}{0.9234} \approx 0.134.$
(D)	Incorrect. This is approximately equal to dividing the probability that a volunteer selected at random will have worked greater than 90 hours by the probability that a volunteer selected at random will have worked between 85.89 hours and 90 hours.
(E)	Incorrect. This is the probability that a volunteer selected at random will have worked less than 90 hours, not the probability that the volunteer selected will receive the certificate of merit given that the number of hours the volunteer worked is less than $90.$

Question 11

Skill		Learning Objective	Topic
2.A		UNC-1.H	Describing the Distribution of a Quantitative Variable
(A)	Incorrect. One of the three values (60) is an outlier.		
(B)	**Correct**. The interquartile range is $76 - 70 = 6$ for the age-group 40 to 50, and 1.5 times the interquartile range is $(1.5)(6) = 9$. Then $\text{Q1} - 9 = 70 - 9 = 61$, and $\text{Q3} + 9 = 76 + 9 = 85$. Of the numbers 60, 62, and 84, only 60 is less than 61 or greater than 85, so 60 is the only outlier.		
(C)	Incorrect. It is true that the value 60 is an outlier. However, the value 62 is not an outlier because 62 is not less than $\text{Q1} - 1.5(\text{IQR})$ or greater than $\text{Q3} + 1.5(\text{IQR})$.		
(D)	Incorrect. It is true that the value 60 is an outlier. However, the value 84 is not an outlier because 84 is not less than $\text{Q1} - 1.5(\text{IQR})$ or greater than $\text{Q3} + 1.5(\text{IQR})$.		
(E)	Incorrect. Only one of the three values (60) is an outlier.		

Question 12

Skill		Learning Objective	Topic
1.C		DAT-2.C	Random Sampling and Data Collection
(A)	Incorrect. A cluster sample involves dividing a population into smaller subgroups. However, the college administrator did not select a simple random sample of all subgroups (majors), and there is no indication that there is heterogeneity within each subgroup (major).		
(B)	Incorrect. A convenience sample was not selected, because a single easily available group of students was not selected to serve as the sample.		
(C)	Incorrect. A simple random sample was not selected, because students were not selected at random from the entire population of students.		
(D)	**Correct**. The administrator selected a stratified random sample, because all of the students at the college were separated into strata (the majors) and a random sample was selected from each of the strata.		
(E)	Incorrect. A systematic random sample was not selected, because the students were selected at random from the majors; it was not the case that every kth student was selected to be in the sample for some integer k.		

Question 13

Skill	Learning Objective	Topic
4.B	UNC-3.Q	Sampling Distributions for Sample Means

(A)	Incorrect. This sampling distribution has the same shape as the population distribution (left-skewed). Because the sample size is sufficiently large, the sampling distribution of the sample mean should be approximately normal.
(B)	**Correct**. For samples of size $40,$ the sampling distribution of the sample mean should be approximately normal, with a mean equal to $\mu_{\overline{x}} = \mu = 85$ and standard deviation equal to $\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}} = \frac{18}{\sqrt{40}} \approx 2.85.$ This graph appears to be approximately normal, centered at $85,$ and with a standard deviation of approximately $2.85.$
(C)	Incorrect. It is correct that the sampling distribution of the sample mean should be approximately normal with a mean of $85.$ However, the standard deviation of the sampling distribution of the sample mean should be equal to $\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}} = \frac{18}{\sqrt{40}} \approx 2.85,$ and the standard deviation in this graph appears to be much less than $2.85.$
(D)	Incorrect. It is correct that the sampling distribution of the sample mean should be approximately normal. However, the sampling distribution of the sample mean should be centered at the population mean of $85,$ not centered at $66.$
(E)	Incorrect. Because the sample size is sufficiently large, the sampling distribution of the sample mean should be approximately normal, not right-skewed. Also, the sampling distribution of the sample mean should be centered at the population mean of $85,$ not at approximately $35.$

Question 14

Skill	Learning Objective	Topic
4.A	UNC-4.H	Justifying a Claim Based on a Confidence Interval for a Population Proportion

(A)	Incorrect. It is true that the interval will be narrower when the sample proportion is farther from 0.5, but in this instance the sample proportion is closer to 0.5.
(B)	Incorrect. The revised interval will be wider, not narrower, for sample proportion values closer to 0.5.
(C)	Incorrect. It is true that the revised interval will be wider than the original interval, but the reason is not because the sample proportion is farther from 0.5 than the miscalculated proportion is.
(D)	**Correct**. The confidence interval is given by the formula $\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$. When the interval is revised, the value of z^* remains the same since the same confidence level is used, and the value of n remains the same. The original value of $\hat{p}$ was the midpoint of the confidence interval (0.17), but it has now changed to 0.27. The greatest value of $\hat{p}(1-\hat{p})$ will occur when $\hat{p} = 0.5$, and the value will decrease for values closer to 0 or 1. Since $\hat{p} = 0.27$ is closer to 0.5 than $\hat{p} = 0.17$, the revised confidence interval will be wider than the original interval since z^* and n remain the same but $\hat{p}(1-\hat{p})$ will increase.
(E)	Incorrect. The original and revised intervals would have the same width only if the values of $\hat{p}$ were the same, but they are different.

Question 15

Skill	Learning Objective	Topic
4.E	DAT-3.B	Concluding a Test for a Population Proportion

(A)	Incorrect. It is true that the data do not provide convincing statistical evidence, but the p-value is very large, so it is not less than any reasonable significance level.
(B)	**Correct**. The test statistic for testing the hypotheses $H_0\colon p = 0.41$ and $H_a\colon p \neq 0.41$ can be found using $z = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$ or technology. The test statistic has the value -0.083, with the corresponding p-value of approximately 0.934 found using technology. This p-value is greater than any reasonable value for the significance level, so the null hypothesis would not be rejected, and the data do not provide convincing statistical evidence that the proportion of all high school students who would respond they are having a good day is different from 0.41.
(C)	Incorrect. The data do <u>not</u> provide convincing statistical evidence, and the p-value is very large, so it is not less than any reasonable significance level.
(D)	Incorrect. It is true that the p-value is greater than any reasonable significance level, but it is not true that the data provide convincing statistical evidence.
(E)	Incorrect. By itself, the expected value of the number of students who will report having a good day does not provide convincing statistical evidence.

Question 16

Skill	Learning Objective	Topic
1.C	VAR-3.A	Introduction to Experimental Design

(A)	Incorrect. Replication exists because there were 10 members assigned to each exercise type, not because there are four types of exercise.
(B)	Incorrect. Replication exists because there were 10 members assigned to each exercise type, not because the experiment was conducted over a six-week period.
(C)	Incorrect. The response variable is the change in maximal oxygen consumption measured, not the type of exercise.
(D)	**Correct**. The values for the explanatory variable (exercise) are the treatments, and these values are strength training, flexibility training, aerobics, and jogging.
(E)	Incorrect. An experimental unit is the smallest unit to which a treatment is applied. Each of the 40 members who participated is an experimental unit, not the four different types of exercise.

Question 17

Skill	**Learning Objective**	**Topic**
3.B	VAR-5.E	Combining Random Variables

(A)	Incorrect. This value was calculated by using a normal distribution with a correct mean of -15 but using a standard deviation that was incorrectly calculated by subtracting the standard deviations of Sean and Evan.
(B)	Incorrect. This value was calculated by using a normal distribution with a correct mean of -15 but by incorrectly using Evan's standard deviation.
(C)	Incorrect. This value was calculated by using a normal distribution with a correct mean of -15 but using a standard deviation that was incorrectly calculated as $\sqrt{25^2 - 15^2} = 20$.
(D)	**Correct**. Let S and E represent Sean's weekly income and Evan's weekly income, respectively. Because S and E are both approximately normal and independent, the distribution of $S - E$ will be approximately normal with mean $\overline{x}_S - \overline{x}_E = 225 - 240 = -15$ and standard deviation $\sqrt{\sigma_{S-E}^2} = \sqrt{\sigma_S^2 + \sigma_E^2} = \sqrt{25^2 + 15^2} = \sqrt{850}$. The probability that Sean's income is greater than Evan's income is $P(S - E > 0)$ in a normal distribution with mean -15 and standard deviation $\sqrt{850}$, which can be found using technology to be approximately 0.303.
(E)	Incorrect. This value was calculated by using a normal distribution with a correct mean of -15 but using a standard deviation that was incorrectly calculated as $25 + 15 = 40$.

Question 18

Skill	Learning Objective	Topic
3.C	UNC-3.L|UNC-3.K	Sampling Distributions for Sample Proportions

(A)	Incorrect. It is correct that the mean is 0.36 and the standard deviation is 0.076. However, the sampling distribution of the sample proportion is approximately normal because the sample size is large enough.
(B)	Incorrect. The sampling distribution of the sample proportion is approximately normal because the sample size is large enough. Also, the mean and standard deviation are not correct. The mean of the sampling distribution of the sample proportion is given by $\mu_{\hat{p}} = p$, and the standard deviation is given by $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$.
(C)	**Correct**. The sampling distribution of the sample proportion is approximately normal because the sample size is large enough ($np = 40(0.36) = 14.4$ and $n(1-p) = 40(1-0.36) = 25.6$, each of which is greater than 10). The mean of the sampling distribution of $\hat{p}$ is $\mu_{\hat{p}} = p = 0.36$, and the standard deviation of the sampling distribution of $\hat{p}$ is $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{(0.36)(0.64)}{40}} \approx 0.076$.
(D)	Incorrect. It is correct that the sampling distribution is approximately normal and the mean is 0.36. However, the standard deviation is incorrect. The standard deviation of the sampling distribution of the sample proportion is given by $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$.
(E)	Incorrect. It is correct that the sampling distribution is approximately normal and the standard deviation is 0.076. However, the mean is incorrect. The mean of the sampling distribution of the sample proportion is given by $\mu_{\hat{p}} = p$.

Question 19

Skill	Learning Objective	Topic
4.B	VAR-3.E	Inference and Experiments

(A)	Incorrect. The 25 student athletes who received the beetroot juice are the athletes in the treatment group, but the results of the study can be generalized to the population from which the sample was selected.
(B)	Incorrect. The 50 student athletes in the sample are the athletes used in the experiment, but the results of the study can be generalized to the population from which the sample was selected.
(C)	**Correct**. The largest population to which the results can be generalized is the population from which the sample was selected, which is all student athletes at the college.
(D)	Incorrect. The results of the study can only be generalized to the population from which the sample was selected, which only includes student athletes at the college, not other students at the college who are not athletes.
(E)	Incorrect. The results of the study can only be generalized to the population from which the sample was selected, which only includes student athletes at the college and does not include other people who exercise but are not from the college.

Question 20

Skill	Learning Objective	Topic
2.D	UNC-1.P	Representing Two Categorical Variables

(A)	Incorrect. Association cannot be determined from the bar graph.
(B)	Incorrect. Association cannot be determined from the bar graph.
(C)	Incorrect. The graph shows the percents of returned surveys, but the numbers cannot be determined unless the total number of surveys is known.
(D)	Incorrect. Symmetric and skewed results have no meaning in the context of the bar graph.
(E)	**Correct**. According to the graph, the rate of return for the Dining Hall delivery method was approximately 33 percent, for the Psychology delivery method was approximately 48 percent, and for the In Class delivery method was approximately 58 percent. The In Class delivery method had the greatest rate of return, and the Dining Hall delivery method had the least rate of return.

Question 21

Skill		Learning Objective	Topic
2.A		DAT-1.G	Least Squares Regression
(A)	Incorrect. This incorrectly describes the meaning of the correlation coefficient r; the correlation coefficient is a measure of the strength of the linear association between age and height and does not give the relationship between an individual age and height.		
(B)	Incorrect. The correlation coefficient r is not equal to the slope of the regression line; the correlation coefficient is a measure of the strength of the linear association between age and height.		
(C)	**Correct**. The coefficient of determination, r^2, is the proportion of the variation in height that is explained by the least-squares regression line. The value of the coefficient of determination is $r^2 = (0.8)^2 = 0.64$, so the proportion of the variation in height that is explained by a regression on age is 0.64.		
(D)	Incorrect. The correlation coefficient r does not give a probability of predicting the height; the correlation coefficient is a measure of the strength of the linear association between age and height.		
(E)	Incorrect. The square of the correlation coefficient, r^2, does not give a probability of predicting the height; the coefficient of determination $\left(r^2\right)$ is the proportion of the variation in the response variable explained by the least-squares regression line.		

Question 22

Skill		Learning Objective	Topic
3.C		UNC-3.R|UNC-3.Q	Sampling Distributions for Sample Means
(A)	Incorrect. It is correct that the sampling distribution of the sample mean is approximately normal and that the mean is 11.4. However, the standard deviation is incorrect. The standard deviation is given by the formula $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$.		
(B)	**Correct**. The distribution of wait times is approximately normal because the sample size of 84 is greater than 30. The mean of the sampling distribution of the sample mean is $\mu_{\bar{x}} = \mu = 11.4$, and the standard deviation of the sampling distribution of the sample mean is $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{2.6}{\sqrt{84}}$.		
(C)	Incorrect. It is correct that the sampling distribution of the sample mean is approximately normal. However, the value of 12.0 is the value for which the manager wishes to calculate a probability; it is not equal to the mean of the sampling distribution. The mean of the sampling distribution of the sample mean is $\mu_{\bar{x}} = \mu$. Also, the value of 2.6 is not the correct standard deviation for the sampling distribution of the sample mean. The standard deviation of the sampling distribution is given by $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$.		
(D)	Incorrect. The distribution of the sample mean is not binomial. The mean and standard deviation are not correct since they are calculated using formulas for the binomial distribution.		
(E)	Incorrect. The distribution of the sample mean is not binomial. The mean and standard deviation are not correct since they are calculated using formulas for the binomial distribution.		

Question 23

Skill	**Learning Objective**	**Topic**
1.E	VAR-7.B	Setting Up a Test for a Population Mean

(A)	Incorrect. The safety officers want to investigate whether there is a mean difference in the number of cars, not a difference between proportions.
(B)	Incorrect. A two-sample z -test for a difference between means is not appropriate because the days on which the number of cars were recorded are not independent. The numbers were recorded on the same days for each school.
(C)	**Correct**. The cars in the investigation are matched by day; the number of cars were recorded for the same day at each school. Because the measurements taken at each school were matched by day and the safety officers want to investigate whether there is an average difference for the 15 differences calculated from the matched pairs, the appropriate test is a matched-pairs t -test for a mean difference.
(D)	Incorrect. A chi-square test is not appropriate because the data is quantitative, not qualitative.
(E)	Incorrect. A chi-square test is not appropriate because the data is quantitative, not qualitative.

Question 24

Skill	**Learning Objective**	**Topic**
2.C	UNC-1.J	Summary Statistics for a Quantitative Variable

(A)	Incorrect. The interquartile range represents the middle 50 percent of the data. There is no interval of width 2 that contains 50 percent of the data values.
(B)	**Correct**. The first quartile, $\text{Q1},$ is the value that has 25 percent of the data values at or below it, so $\text{Q1} = 66.$ The third quartile, $\text{Q3},$ is the value that has 25 percent of the data values at or above it, so $\text{Q3} = 71.$ The interquartile range is $\text{Q3} - \text{Q1} = 71 - 66 = 5.$
(C)	Incorrect. The interquartile range represents the middle 50 percent of the data. There is no interval of length 9 such that 25 percent of the data values are less than the left endpoint and 25 percent of the data values are greater than the right endpoint.
(D)	Incorrect. The interquartile range represents the middle 50 percent of the data. There is no interval of length 12 such that 25 percent of the data values are less than the left endpoint and 25 percent of the data values are greater than the right endpoint.
(E)	Incorrect. The interquartile range represents the middle 50 percent of the data. There is no interval of length 15 such that 25 percent of the data values are less than the left endpoint and 25 percent of the data values are greater than the right endpoint.

Question 25

Skill	Learning Objective	Topic
4.E	DAT-3.F	Carrying Out a Test for a Population Mean

(A)	**Correct**. The hypotheses tested are $H_0: \mu = 13$ versus $H_a: \mu < 13$. The test statistic is equal to $t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{12.5 - 13}{\frac{6.5}{\sqrt{400}}} \approx -1.54$, with the number of degrees of freedom equal to $n - 1 = 400 - 1 = 399$. The p-value is 0.0624, found using technology. Since the p-value is greater than the value of alpha $(0.0624 > 0.05)$, the null hypothesis is not rejected and there is not convincing statistical evidence to conclude that the average number of hours worked per week at part-time jobs decreased after the salary increase.
(B)	Incorrect. It is correct that there is not convincing statistical evidence to conclude that the average number of hours worked per week at part-time jobs decreased after the salary increase. However, the p-value of the appropriate test is not less than 0.05.
(C)	Incorrect. It is incorrect that there is convincing statistical evidence, but it is correct that the p-value of the appropriate test is greater than 0.05.
(D)	Incorrect. It is incorrect that there is convincing statistical evidence, and it is also incorrect that the p-value of the appropriate test is less than 0.05.
(E)	Incorrect. There is enough information to conduct the appropriate hypothesis test and to make a conclusion.

Question 26

Skill		Learning Objective	Topic
3.B		VAR-5.C	Mean and Standard Deviation of Random Variables
(A)	Incorrect. This is the probability that there are 2 people in a passenger car.		
(B)	Incorrect. This is the probability that there is 1 person in a passenger car.		
(C)	**Correct**. The mean number of people in passenger cars is $1(0.56)+2(0.28)+3(0.08)+4(0.06)+5(0.02)=1.7$.		
(D)	Incorrect. The department will base their recommendation on this number of people.		
(E)	Incorrect. This is the mean of the numbers of people, $\frac{1+2+3+4+5}{5}=3$.		

Question 27

Skill		Learning Objective	Topic
1.B		VAR-3.B	Introduction to Experimental Design
(A)	Incorrect. It is not a requirement that the number of subjects in each block in a randomized block design be different. The number of subjects in each block can be equal or different.		
(B)	**Correct**. A feature of a well-designed experiment is randomization, which reduces the chance of bias in experimental groups. Randomization can be achieved in an experiment by randomly assigning treatments to subjects within each block.		
(C)	Incorrect. Blocking by age-group does not mean that there cannot be a control group.		
(D)	Incorrect. There is no matching between groups in this experiment. The subjects in one group and the subjects in the other group are different and not paired in any way.		
(E)	Incorrect. In a randomized block design, subjects within each of the blocks are randomly assigned to the two treatments.		

Question 28

Skill		Learning Objective	Topic
3.A		UNC-3.E	The Geometric Distribution
(A)	Incorrect. The value 0.1406 represents the probability that a color other than blue lands faceup on the first toss, followed by a color other than blue on the second toss, followed by a blue on the third toss, which is not equal to the probability that the player will toss the die at least 2 times before blue lands faceup.		
(B)	Incorrect. The value 0.4219 represents the probability that a color other than blue lands faceup 3 times when the die is tossed 3 times, which is not equal to the probability that the player will toss the die at least 2 times before blue lands faceup.		
(C)	Incorrect. The value 0.4375 represents the probability that a player will toss the die fewer than 2 times before blue lands faceup, which is not equal to the probability that the player will toss the die at least 2 times before blue lands faceup.		
(D)	**Correct**. Let B represent the number of tosses until a blue lands faceup. The random variable B follows a geometric distribution with $p = 0.25$. The probability that a player will toss the die at least 2 times before blue lands faceup is $P(B \geq 3) = 1 - P(B < 3) = 1 - [P(B = 2) + P(B = 1)] = 1 - [0.25 + (0.25)(0.75)]$.		
(E)	Incorrect. The value 0.5781 represents the probability that a player will toss the die fewer than 3 times before blue lands faceup, which is not equal to the probability that the player will toss the die at least 2 times before blue lands faceup.		

Question 29

Skill	Learning Objective	Topic
1.B	UNC-5.A	Potential Errors When Performing Tests
(A)	Incorrect. Failing to reject the null hypothesis is a correct decision, not an error, when the null hypothesis is true.	
(B)	**Correct**. A Type II error occurs when the null hypothesis is not rejected but it should have been rejected. Not rejecting the null hypothesis means that a conclusion is reached where there is not enough statistical evidence to conclude that the population mean is greater than 64, but in fact the population mean is greater than 64.	
(C)	Incorrect. Rejecting the null hypothesis when the null hypothesis is true is a Type I error, not a Type II error.	
(D)	Incorrect. Rejecting the null hypothesis when the population mean is greater than 64 is a correct decision, not an error.	
(E)	Incorrect. Failing to reject the null hypothesis when the p-value is less than the signficance level is an incorrect decision, but it is neither a Type I nor Type II error.	

Question 30

Skill	Learning Objective	Topic
4.B	UNC-4.AA	Justifying a Claim About the Difference of Two Means Based on a Confidence Interval

(A)	Incorrect. The values in the interval are all negative, which is necessary if mango has the greater sample mean rating, but the difference in means must be between -4 and 0, and these values do not meet that condition.
(B)	**Correct**. If there was a statistically significant difference in mean flavor rating, with mango having the greater sample mean rating, then the difference in means (cotton candy minus mango) must be negative. Also, the difference in means must be between -4 and 0 because the ratings for each flavor were between 1 and 5 and mango had the greater sample mean rating. Of the intervals listed, only $(-2.1, -1.3)$ has values that are all negative between -4 and 0.
(C)	Incorrect. The interval represents the set of plausible values for the difference in population means. Because the interval contains negative values, 0, and positive values, it is plausible that the cotton candy mean is greater than the mango mean. It is also plausible that there is no difference in population means (as indicated by 0 in the interval).
(D)	Incorrect. The interval represents the set of plausible values for the difference in population means. This interval provides evidence that the cotton candy mean is greater than the mango mean because all values in the interval are positive.
(E)	Incorrect. The interval represents the set of plausible values for the difference in population means. This interval provides evidence that the cotton candy mean is greater than the mango mean because all values in the interval are positive. Also, it is not possible to construct this interval from the data because the flavors are rated on a scale of 1 to 5.

Question 31

Skill	Learning Objective	Topic
3.B	UNC-3.K	Sampling Distributions for Sample Proportions

(A)	Incorrect. There is no variance associated with a single sample proportion.
(B)	Incorrect. There is no variance associated with a single population proportion.
(C)	**Correct**. The variance of the sampling distribution of the sample proportion is given by $\sigma_{\hat{p}}^2 = \frac{p(1-p)}{n}$. If the value of n is decreased, the value of the fraction will increase. Therefore, the variance of the sampling distribution of the sample proportion will increase.
(D)	Incorrect. As sample size decreases, the variance of the sampling distribution of the sample proportion will increase, not decrease.
(E)	Incorrect. The variance of the sampling distribution of the sample proportion will change as the value of n changes in the formula $\sigma_{\hat{p}}^2 = \frac{p(1-p)}{n}$.

Question 32

Skill	**Learning Objective**	**Topic**
3.D	UNC-4.C	Constructing a Confidence Interval for a Population Proportion

(A)	Incorrect. The z^* value used in the confidence interval formula is for a 95 percent confidence interval, not a 90 percent confidence interval. Also, the square root should contain the entire fraction, not just the denominator of the fraction. The correct confidence interval formula is given by $\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$.
(B)	Incorrect. The square root should contain the entire fraction, not just the denominator of the fraction. The correct confidence interval formula is given by $\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$.
(C)	Incorrect. The z^* value used in the confidence interval formula is for a 99 percent confidence interval, not a 90 percent confidence interval.
(D)	Incorrect. The z^* value used in the confidence interval formula is for a 95 percent confidence interval, not a 90 percent confidence interval.
(E)	**Correct**. The formula $\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ gives a confidence interval for one-sample proportion. Technology can be used to find critical value z^* for a 90 percent confidence interval. Substituting the values $\hat{p} = 0.32$, $z^* = 1.645$, and $n = 1{,}005$ into the confidence interval formulas yields $\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.32 \pm 1.645\sqrt{\frac{(0.32)(1-0.32)}{1{,}005}} = 0.32 \pm 1.645\sqrt{\frac{(0.32)(0.68)}{1{,}005}}$.

Question 33

Skill	**Learning Objective**	**Topic**
4.B	UNC-4.AF	Confidence Intervals for the Slope of a Regression Model

(A)	Incorrect. This interval was obtained by incorrectly using the t-value in the computer output (3.27) in the formula for the confidence interval, which is not the t-value for the confidence interval. The t-value for a confidence interval for the slope is found in a t-table, or using technology for the t-distribution with 29 degrees of freedom.
(B)	**Correct**. The interval estimate for the slope of a regression model is given by the formula $b \pm t^{*}(\text{SE}_b)$, where b is the slope of the line of best fit, and SE_b is the standard error for the slope of the regression line. The value of b is the estimate of the diameter in the computer output (1.054), and the value of SE_b is the standard error of the diameter in the computer output (0.322). The value of t^{*} for a 95 percent confidence interval is found using technology to be 2.045, with $n - 2 = 31 - 2 = 29$ degrees of freedom. The confidence interval is thus $1.054 \pm 2.045(0.322)$, which yields the confidence interval $(0.396, 1.712)$.
(C)	Incorrect. This confidence interval was calculated using correct values for b and SE_b in the confidence interval formula $b \pm t^{*}(\text{SE}_b)$, but incorrectly used the z^{*} value for a 95 percent interval, not a t^{*} value with $n - 2 = 31 - 2 = 29$ degrees of freedom.
(D)	Incorrect. This confidence interval used the incorrect formula $b \pm (\text{SE}_b)$, which omits the required t^{*} value. The correct formula is $b \pm t^{*}(\text{SE}_b)$.
(E)	Incorrect. This confidence interval used the values for the estimate and standard error for the intercept in the formula but should have used the values of the estimate and standard error for the diameter in the formula. The correct value of t^{*} was used in the formula.

Question 34

Skill	Learning Objective	Topic
4.E	DAT-3.D	Carrying Out a Test for the Difference of Two Population Proportions

(A)	Incorrect. Randomization was used in the study to randomly assign treatments to the volunteer subjects, so a conclusion can be made.
(B)	Incorrect. The p-value of 0.1645 for the hypothesis test is greater than $0.01,$ so the null hypothesis is not rejected and there is insufficient statistical evidence to conclude that the proportion of people who would be classified as normal after taking cinnamon is greater than the proportion who would be classified as normal after not taking cinnamon.
(C)	Incorrect. The p-value of 0.1645 for the hypothesis test is greater than either 0.01 or $0.05,$ so the null hypothesis is not rejected and there is insufficient statistical evidence to conclude that the proportion of people who would be classified as normal after taking cinnamon is greater than the proportion who would be classified as normal after not taking cinnamon.
(D)	Incorrect. The p-value of 0.1645 for the hypothesis test is greater than either 0.10 or $0.05,$ so the null hypothesis is not rejected and there is insufficient statistical evidence to conclude that the proportion of people who would be classified as normal after taking cinnamon is greater than the proportion who would be classified as normal after not taking cinnamon.
(E)	**Correct**. A two-sample z-test for a difference in population proportions can be conducted to test the hypothesis $\mathrm{H}_0\colon p_1 - p_2 = 0$ versus $\mathrm{H}_\mathrm{a}\colon p_1 - p_2 > 0,$ where the subscript 1 represents the cinnamon group and the subscript 2 represents the placebo group. The combined (or pooled) proportion needed for the test is given by $\hat{p}_c = \dfrac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2} = \dfrac{40\left(\frac{14}{40}\right) + 40\left(\frac{10}{40}\right)}{40 + 40} = 0.3.$ The test statistic is equal to $z = \dfrac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\hat{p}_c(1 - \hat{p}_c)}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \dfrac{\frac{14}{40} - \frac{10}{40}}{\sqrt{\frac{3}{10}\left(1 - \frac{3}{10}\right)}\sqrt{\frac{1}{40} + \frac{1}{40}}} \approx 0.976.$ The corresponding p-value, found using technology, is approximately $0.1645,$ which is very large, so there is not convincing statistical evidence at any reasonable significance level.

Question 35

Skill	Learning Objective	Topic
4.C	VAR-7.L	Setting Up a Test for the Slope of a Regression Model

(A)	Incorrect. A residual plot does not indicate if the errors from a sample are independent. To check for independence, data should be collected using a random sample or a randomized experiment, and when a sample is selected without replacement, the sample size must be less than or equal to 10 percent of the population size.
(B)	Incorrect. It is true that the sum of the residuals is 0, but this is not a condition for the test which must be checked.
(C)	Incorrect. It is true that the expected value of the errors is 0, but this is not a condition for the test which must be checked.
(D)	Incorrect. This is a condition for the test to be checked. However, the residual plot is not the most appropriate display to check this condition. A scatterplot of the explanatory variable and response variable is more appropriate to check this condition.
(E)	**Correct**. To test the claim that the maximum height and the maximum speed are linearly related, one of the conditions that must be satisfied is that the residuals must have constant error variance. The displayed residuals are not evenly spread around the horizontal line at 0 since the residual points are closer to the line for heights below 125 and further from the line for heights greater than 125. Thus the requirement of constant error variance for all values of the explanatory variable has not been satisfied.

Question 36

Skill	**Learning Objective**	**Topic**
4.B	DAT-3.A	Interpreting P-Values

(A)	Incorrect. This is the probability of obtaining a sample statistic that is not as extreme as the one observed under the assumption that the null hypothesis in the original set of hypotheses is true. However, it cannot be a p-value, since a p-value is the probability of obtaining a test statistic that is as extreme or more extreme than the test statistic observed under the assumption that the null hypothesis is true.
(B)	Incorrect. The value $2(0.0627)$ is the area in the tails of a two-tailed test corresponding to an alternative hypothesis containing a hypothesized value different from $38.$ Therefore, the value $1-2(0.0627)$ is not equal to the p-value.
(C)	Incorrect. The new test is left tailed, and the value $\frac{1}{2}(0.0627)$ is the area in the left tail. The value $1-\frac{1}{2}(0.0627)$ is the probability of obtaining a sample statistic that is not as extreme as the one observed under the assumption that the null hypothesis in the original set of hypotheses is true, so does not meet the definition of a p-value.
(D)	Incorrect. The new alternative hypothesis corresponds to a left-tailed test, so the area in the left tail should be half of what the area in the two tails was, not twice that area.
(E)	**Correct**. A p-value is the probability of obtaining a test statistic as extreme or more extreme than the test statistic observed under the assumption that the null hypothesis is true. The original set of hypotheses indicates that a two-tailed test is to be conducted, which means that the p-value comprises the sum of the area in the right tail and the area in the left tail. Also, the areas in the tails are equal. If the alternative hypothesis is changed so that the test is left tailed, then the p-value is halved to find the area in only the left tail. Thus the p value would have been $\frac{1}{2}(0.0627)$.

Question 37

Skill	Learning Objective	Topic
3.B	VAR-5.E	Combining Random Variables

(A)	Incorrect. It is true that the mean is 34 seconds. It is not true, however, that the variables are independent, since X and Y represent the running times before and after training for the same student, and it is not true that the standard deviation is 10 seconds.
(B)	Incorrect. It is true that the mean is 34 seconds. It is not true, however, that the variables are independent, since X and Y represent the running times before and after training for the same student, and it is not true that the standard deviation is 50 seconds.
(C)	Incorrect. It is true that the variables X and Y are not independent, since X and Y represent the running times before and after training for the same student. There is, however, enough information to calculate the mean, but there is not enough information provided to calculate the standard deviation.
(D)	**Correct**. The random variables X and Y represent the running times before and after training for the same student, so the variables are dependent, not independent. The mean of $X - Y$ is $\mu_{X-Y} = \mu_X - \mu_Y = 402 - 368 = 34$ seconds. If X and Y are independent, the variance σ^2_{X-Y} of $X - Y$ is equal to $\sigma^2_X + \sigma^2_Y$. Since X and Y are not independent, the variance and hence the standard deviation cannot be determined with the given information.
(E)	Incorrect. It is true that the variables X and Y are not independent, since X and Y represent the running times before and after training for the same student, and it is true that there is not enough information to calculate the standard deviation. There is, however, enough information to calculate the mean.

Question 38

Skill	**Learning Objective**	**Topic**
3.D	UNC-4.K	Confidence Intervals for the Difference of Two Proportions

(A)	**Correct**. Let the subscript 1 denote adults who are not scientists, and the subscript 2 denote adults who are scientists. Then $n_1 = 2{,}002$, $n_2 = 3{,}748$, $p_1 = 0.37$, and $p_2 = 0.88$. The standard error is equal to $\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = \sqrt{\frac{(0.37)(0.63)}{2{,}002}+\frac{(0.88)(0.12)}{3{,}748}}$.
(B)	Incorrect. In this response, the fractions are subtracted, instead of added, in the formula for standard deviation. The standard error is given by $\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$.
(C)	Incorrect. The two fractions should be added under one square root, not added after the square root is applied to each fraction. The standard error is given by $\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$.
(D)	Incorrect. The two fractions should be added under one square root, not added after the square root is applied to each fraction. Also, the pooled proportion is incorrectly used for $\hat{p}_1$ and $\hat{p}_2$. The standard error is given by $\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$.
(E)	Incorrect. The correct values for the sample sizes and proportions were used. A square root of the sum of two fractions should have been used, however, but the square root was applied only to the denominators. The standard error is given by $\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$.

Question 39

Skill	Learning Objective	Topic
3.E	VAR-8.L	Carrying Out a Chi-Square Test for Homogeneity or Independence

(A)	Incorrect. This incorrectly used $\text{observed count} - \text{expected count}$ in the calculation, but should have used $(\text{observed count} - \text{expected count})^2$. The correct formula is $\frac{(\text{observed count} - \text{expected count})^2}{\text{expected count}}$.
(B)	Incorrect. This divided by the observed count in the calculation, but should have divided by the expected count. The correct formula is $\frac{(\text{observed count} - \text{expected count})^2}{\text{expected count}}$.
(C)	**Correct**. The chi-square test statistic is calculated by summing the values $\frac{(\text{observed count} - \text{expected count})^2}{\text{expected count}}$. The expected count is found by $\frac{(\text{row total})(\text{column total})}{\text{table total}} = \frac{(1{,}000)(60)}{2{,}000} = 30.$ The contribution to the test statistic is equal to $\frac{(\text{observed count} - \text{expected count})^2}{\text{expected count}} = \frac{(45-30)^2}{30} = 7.5.$
(D)	Incorrect. This is the expected count, not the contribution to the chi-square test statistic.
(E)	Incorrect. This is the count of men who considered business networking important, but it is not the contribution to the chi-square test statistic.

Question 40

Skill		Learning Objective	Topic
4.B		UNC-4.S	Justifying a Claim About a Population Mean Based on a Confidence Interval
(A)	Incorrect. The percent is how much confidence exists that the interval has captured the population mean; it is not about the percentage of individual observations in the population that fall within the interval.		
(B)	Incorrect. Once the interval is constructed, the interpretation of the confidence interval should not be a statement about probability. Once the sample has been selected and the interval constructed, the unknown population mean was either captured by the interval (probability equal to 1) or not (probability equal to 0).		
(C)	Incorrect. Different samples can yield different results. The interval is a statement about how confident we are that we have captured the population parameter, not any possible sample proportion.		
(D)	Incorrect. The interval is used to estimate the unknown population mean, not the sample mean. The sample mean is not estimated. It is used to create the interval and will always be at the midpoint of the interval.		
(E)	**Correct**. The percent is how much confidence exists that the interval has captured the population mean.		

2019 AP Statistics
Question Descriptors and Performance Data

Multiple-Choice Questions

Question	Skill	Learning Objective	Topic	Key	% Correct
1	2.A	UNC-1.H	Describing the Distribution of a Quantitative Variable	C	61
2	3.A	VAR-4.D	Conditional Probability	E	72
3	2.A	DAT-1.F	Residuals	C	63
4	2.A	UNC-1.H	Describing the Distribution of a Quantitative Variable	A	79
5	2.C	UNC-1.Q	Statistics for Two Categorical Variables	A	78
6	3.A	VAR-2.B	The Normal Distribution	C	65
7	2.D	UNC-1.N	Comparing Distributions of a Quantitative Variable	C	92
8	1.C	DAT-2.C	Random Sampling and Data Collection	E	73
9	3.A	VAR-2.B	The Normal Distribution	B	62
10	3.A	VAR-6.B	The Normal Distribution, Revisited	C	21
11	2.A	UNC-1.H	Describing the Distribution of a Quantitative Variable	B	60
12	1.C	DAT-2.C	Random Sampling and Data Collection	D	76
13	4.B	UNC-3.Q	Sampling Distributions for Sample Means	B	58
14	4.A	UNC-4.H	Justifying a Claim Based on a Confidence Interval for a Population Proportion	D	35
15	4.E	DAT-3.B	Concluding a Test for a Population Proportion	B	62
16	1.C	VAR-3.A	Introduction to Experimental Design	D	76
17	3.B	VAR-5.E	Combining Random Variables	D	33
18	3.C	UNC-3.L|UNC-3.K	Sampling Distributions for Sample Proportions	C	68
19	4.B	VAR-3.E	Inference and Experiments	C	76
20	2.D	UNC-1.P	Representing Two Categorical Variables	E	86
21	2.A	DAT-1.G	Least Squares Regression	C	42
22	3.C	UNC-3.R|UNC-3.Q	Sampling Distributions for Sample Means	B	74
23	1.E	VAR-7.B	Setting Up a Test for a Population Mean	C	31
24	2.C	UNC-1.J	Summary Statistics for a Quantitative Variable	B	69
25	4.E	DAT-3.F	Carrying Out a Test for a Population Mean	A	56
26	3.B	VAR-5.C	Mean and Standard Deviation of Random Variables	C	79
27	1.B	VAR-3.B	Introduction to Experimental Design	B	77
28	3.A	UNC-3.E	The Geometric Distribution	D	45
29	1.B	UNC-5.A	Potential Errors When Performing Tests	B	67
30	4.B	UNC-4.AA	Justifying a Claim About the Difference of Two Means Based on a Confidence Interval	B	52
31	3.B	UNC-3.K	Sampling Distributions for Sample Proportions	C	41

2019 AP Statistics
Question Descriptors and Performance Data

Question	Skill	Learning Objective	Topic	Key	% Correct
32	3.D	UNC-4.C	Constructing a Confidence Interval for a Population Proportion	E	74
33	4.B	UNC-4.AF	Confidence Intervals for the Slope of a Regression Model	B	32
34	4.E	DAT-3.D	Carrying Out a Test for the Difference of Two Population Proportions	E	43
35	4.C	VAR-7.L	Setting Up a Test for the Slope of a Regression Model	E	25
36	4.B	DAT-3.A	Interpreting P-Values	E	55
37	3.B	VAR-5.E	Combining Random Variables	D	28
38	3.D	UNC-4.K	Confidence Intervals for the Difference of Two Proportions	A	77
39	3.E	VAR-8.L	Carrying Out a Chi-Square Test for Homogeneity or Independence	C	40
40	4.B	UNC-4.S	Justifying a Claim About a Population Mean Based on a Confidence Interval	E	66

Free-Response Questions

Question	Skill	Learning Objective	Topic	Mean Score
1	2.A|4.B	UNC-1.K|UNC-1.H|UNC-1.M	1.7|1.6|1.8	2.25
2	1.D|3.D|4.B|4.D	UNC-4.AC|UNC-4.AF|UNC-4.AG|UNC-4.AH	9.2|9.3	1.2
3	2.A|2.D	UNC-1.N|UNC-1.M	1.9|1.8	2.19
4	1.B|1.C	VAR-3.D|VAR-3.B	3.6|3.5	1.47
5	3.A|3.C	UNC-3.B|VAR-4.E|UNC-3.L	4.10|4.6|5.5	0.97
6	2.A|2.B|4.B|4.C	VAR-7.H|UNC-1.M	7.8|1.8	1.93