Important Instructions for this Test: Please pay close attention to and carefully follow all rounding instructions. Round any intermediate steps as indicated or as necessary to make the final answer as accurate as possible.
Good luck, and as always: "NOTA" stands for "None of These Answers is correct."

You will need to use the following information to answer every question on this test:
Señor Nieve recently gave a test to his AP Statistics students in a single class. The scores on the test are displayed in the table below. They follow a standard 100-point scale of 0 to 100 points possible and are sorted from highest to lowest along with the observed Z-score of each test score, which are calculated using the sample mean and the sample standard deviation of the set of test scores in the table. The expected Z-score of each test score is also provided. These are based upon the assumption that the set of observed test scores were drawn from a theoretical population distribution of all possible test scores that the students could have earned that is normally distributed.

Rank	Test Score	Z-Score Observed	Z-Score Expected
1	100	1.51125719	1.73166440
2	96	1.07946942	1.15034938
3	96	1.07946942	0.81221780
4	92	0.64768165	0.54852228
5	88	0.21589388	0.31863936
6	88	0.21589388	0.10463346
7	84	-0.21589388	-0.10463346
8	84	-0.21589388	-0.31863936
9	80	-0.64768165	-0.54852228
10	80	-0.64768165	-0.81221780
11	76	-1.07946942	-1.15034938
12	68	-1.94304496	-1.73166440

1. Compute the sample mean, the sample median, and the sample standard deviation of the set of 12 test scores. What is the sum of these three values rounded to four decimal places?

A: 94.8694 B: 95.2638 C: 180.8694 D: 181.2638 E: NOTA

2. Which of the following statements is / are true?

I. The mean of the twelve z-scores in the "Z-Score Observed" column is equal to the mean of the twelve z-scores in the "Z-Score Expected" column.
II. The sample standard deviation of the twelve z-scores in the "Z-Score Observed" column is equal to the sample standard deviation of the twelve z-scores in the "Z-Score Expected" column.

A: Statement I only B: Statement II only C: Both are true D: Neither are true E: NOTA

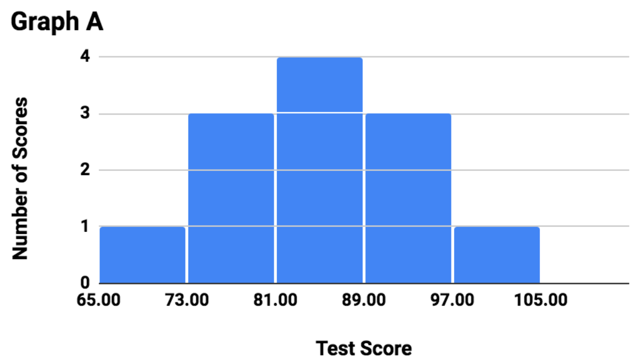
3. Obviously, there are no high outliers in the sample data set of 12 test scores in the table above; however, using the 1.5(IQR) rule for outliers, what is the absolute value of the difference between the current lowest test score in the table above and the cutoff score for determining a low outlier under the assumption that the true theoretical population distribution of all possible test scores on this test follows a normal distribution with a mean of 80 and a standard deviation of 8? Round all steps and the final answer to the nearest tenth.

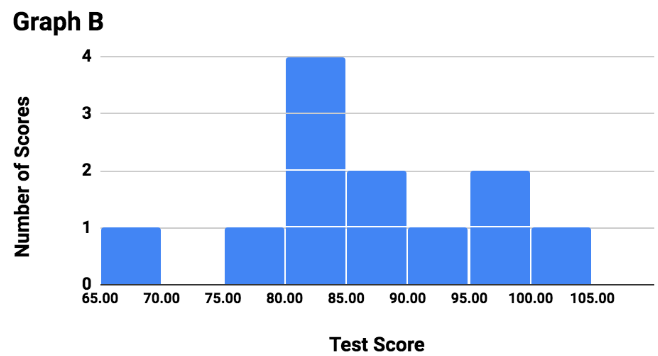
A: 16.2 B: 9.6 C: 21.0 D: 9.0 E: NOTA

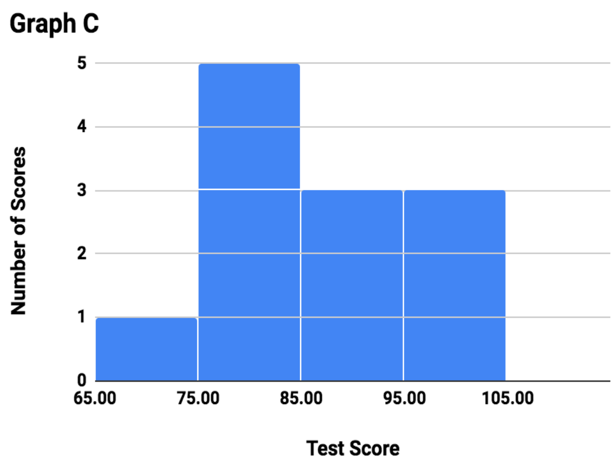
4. What is the coefficient of determination between the set of z-scores in the "Z-Score Observed" column in the table and the set of z-scores in the "Z-Score Expected" column in the table rounded to four decimal places?

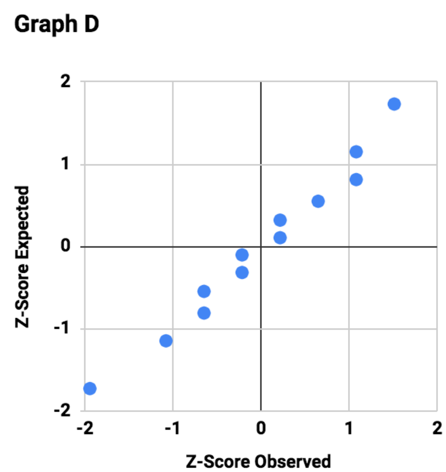
A: 1.0000 B: 0.9878 C: 0.9783 D: 0.9757 E: NOTA

Use the following additional information to answer the next few Questions:
Each of the four graphs below may (or may not) have been created from the various sets of scores in the table on the previous page – that is part of what you need to determine. Examine each graph carefully and answer the questions that follow.









5. Which of the four graphs above is not constructed from values in the table on the previous page; but rather, it is constructed from some other set of values that are not actually in the table on the previous page?

A: Graph A B: Graph B C: Graph C D: Graph D E: NOTA

6. Which of the four graphs above is a standardized normal quantile plot (also known as a "Q-Q plot")?

A: Graph A B: Graph B C: Graph C D: Graph D E: NOTA

7. Is it reasonable for Señor Nieve to conclude that this data set of 12 test scores was indeed drawn from a theoretical population distribution of all possible test scores that is normally distributed? Why or why not?

A: No, because although Graph A shows evidence that the distribution of the data is symmetric, it is not sufficiently close enough to being approximately normal.
B: No, because Graph B shows evidence that the distribution of the data is skewed to the right and it even has a low outlier and so the data could not possibly have come from a normal distribution.
C: No, because Graph C shows evidence that the distribution of the data is slightly skewed to the right and is not at all bell-shaped and so the data could not possibly have come from a normal distribution.
D: Yes, because Graph D shows a strong linear pattern evenly scattered along the line y = x.
E: NOTA

8. Señor Nieve is also an amateur psychometrician and wishes to assess the validity of this particular test based upon the 12 resulting test scores in the table. Scores on the questions from the standardized test bank of questions from which the actual questions on this particular test were drawn is intended to follow a normal distribution with a mean of 80. However, it appears that this set of students may have outscored that intended mean. Señor Nieve is willing to view the set of 12 test scores in the table as a SRS from the theoretical population distribution of all possible test scores this group of students could have earned and that all students' scores are independent of each other. Perform the appropriate statistical test on the data to determine if there is evidence that this group of students has a higher mean score on this test than the intended theoretical population mean of 80. What is the sum of the test statistic, the p-value, and the degrees of freedom of the test rounded to <u>four decimal places</u>? NOTE: The standard deviation of the population distribution of all possible test scores these 12 students could have earned is <u>not</u> assumed to be known.

A: 14.2668 B: 14.2900 C: 13.2668 D: 13.2886 E: NOTA

9. Señor Nieve is also wondering if the theoretical population distribution of all possible test scores students could earn on this test has a larger standard deviation than the one of the intended distribution in the test bank of possible scores, which is 8. Assessing this requires a chi-square hypothesis test for the variance (or equivalently the standard deviation) of a normally distributed population. The test statistic of this test follows a chi-square distribution with n – 1 degrees of freedom and is calculated according to the following formula: $\chi^2_{n-1} = \frac{(n-1)s^2}{\sigma^2}$; where s^2 is the observed sample variance and σ^2 is the assumed population variance in the null hypothesis. Perform this test using the given formula and determine if these data provide evidence that the theoretical population distribution of all possible test scores these 12 students could have earned on this test has a larger standard deviation than the one of the intended distribution in the test bank of possible scores. You may assume all conditions for the test are met. What is the p-value of the test rounded to <u>four decimal places</u>?

A: 0.1942 B: 0.2554 C: 0.3108 D: 0.3884 E: NOTA

10. Given the results of the two hypothesis tests Señor Nieve conducted in the previous two questions, the data set of 12 test scores leads him to which of the following conclusions regarding the mean and the standard deviation of the theoretical population distribution of all possible test scores the students could have earned? Make your decision about each test independent of the other at the 5% level of significance.

A: The data provide statistically significant evidence that the theoretical population distribution has both a mean that is greater than 80 and a standard deviation that is greater than 8.

B: The data provide statistically significant evidence that the theoretical population distribution has a mean that is greater than 80; but they do not provide statistically significant evidence that the standard deviation is greater than 8.

C: The data do not provide statistically significant evidence that the theoretical population distribution has a mean that is greater than 80; but they do provide statistically significant evidence that the standard deviation is greater than 8.

D: The data do not provide statistically significant evidence that the theoretical population distribution has a mean that is greater than 80; nor they do provide statistically significant evidence that the standard deviation is greater than 8.

E: NOTA

11. Initially, the set of test scores in the table on the previous page consisted of 12 students' scores. However, there were actually 14 students in the class. One student made up the test and earned a score of 84 while the other student earned a score of 60 when he made up the test. After these two additional students' scores are added to the data set, all the summary statistics were recomputed. Which of the following statistics is completely unaffected by the addition of these two new scores to the data set?

A: the mean B: the median C: the standard deviation D: the IQR E: NOTA

12. Believe it or not, definitively assessing the "skewness" of a data set is not as easy as many people think and many misconceptions abound, especially when the size of the data set is relatively small; such as 12 as we have here in the original set of test scores for the students in Señor Nieve's class. One formal theoretical measure of the skewness of a probability distribution is known as the "third standardized moment" and is denoted and calculated as follows: $\tilde{\mu}_3 = \frac{\mu_3}{\sigma^3} = \frac{E[(X-\mu)^3]}{(E[(X-\mu)^2])^{3/2}}$. Now, if you look at this formula closely, you will see that it resembles the cube of a z-score. This is no accident and it gives us insight into how this parameter is estimated from sample data, namely: $g_1 = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\bar{X}}{s_x}\right)^3 = \frac{1}{n}\sum_{i=1}^{n} Z_i^3$. This is now literally the mean of the cubed z-scores of the data and is often referred to as "the sample standardized skewness coefficient." However, we are not quite done. We also need to adjust this formula slightly for the sample size in order to give this statistic a very special and important property. We do this for the same reason why we divide by n – 1 instead of by n in the formula for the sample variance so that we can call the sample standardized skewness coefficient, g_1, a(n) _______ estimator of the population standardized skewness coefficient, $\tilde{\mu}_3$. What is the missing word in the blank?

A: resistant B: unbiased C: highly accurate D: low variability E: NOTA

13. The formula for the adjusted sample standardized skewness coefficient defined in the previous question is given by the following formula: $g_1 = \frac{n^2}{(n-1)(n-2)}\left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\bar{X}}{s_x}\right)^3\right] = \frac{n}{(n-1)(n-2)}\left[\sum_{i=1}^{n} Z_i^3\right]$. Compute the adjusted sample standardized skewness coefficient for the original set of 12 test scores in the table on the first page. Round all steps to at least four decimal places and the final answer to the thousandths place.

A: 0.000 B: -0.026 C: -0.242 D: -0.316 E: NOTA

14. Under the assumption that underlying random variable modeling the population from which a random sample of data came is normally distributed, the sampling distribution of the adjusted sample standardized skewness coefficient given in the previous question also follows a normal distribution with a mean of $\mu_{g_1} = E(g_1) = 0$ and a variance of $\sigma_{g_1}^2 = Var(g_1) = \frac{6n(n-1)}{(n-2)(n+1)(n+3)}$. Use this information to construct a 95% confidence interval for the population standardized skewness coefficient using your final rounded result form the previous question as your point estimate of the parameter. Round your final confidence interval limits to three decimal places; and naturally, you may assume all requisite inference assumptions and conditions are met.

A: $-1.275 < \tilde{\mu}_3 < 1.223$
B: $-1.112 < \tilde{\mu}_3 < 0.480$
C: $-1.565 < \tilde{\mu}_3 < 0.933$
D: $-1.491 < \tilde{\mu}_3 < 1.007$
E: NOTA

15. Drawing an inference about the standardized skewness coefficient in the population from which a sample is drawn works essentially the same as it does for any other parameter. Suppose we assume that the population distribution of all possible test scores from which the SRS of 12 test scores in the table on the first page came is normally distributed. Not surprisingly, the population standardized skewness coefficient of a normal distribution is $\tilde{\mu}_3 = 0$, indicating no skewness because a normal distribution is perfectly symmetric. Thus, our null and alternative hypotheses become $H_0: \tilde{\mu}_3 = 0 \ vs.\ H_A: \tilde{\mu}_3 \neq 0$ for a two-tailed test. Use the 95% confidence interval from the previous question to make a decision regarding this set of hypotheses and draw an inference with respect to whether or not we can reject the assumption that this sample of 12 test scores was drawn from a normally distributed population at the 5% level of significance.

A: Since the 95% confidence interval captures 0, we fail to reject the null hypothesis and so we fail to reject the assumption that the set of test scores was drawn from a normally distributed population.
B: Since the 95% confidence interval fails to capture 0, we fail to reject the null hypothesis and so we fail to reject the assumption that the set of test scores was drawn from a normally distributed population.
C: Since the 95% confidence interval captures 0, we reject the null hypothesis and so we reject the assumption that the set of test scores was drawn from a normally distributed population.
D: Since the 95% confidence interval fails to captures 0, we reject the null hypothesis and so we reject the assumption that the set of test scores was drawn from a normally distributed population.
E: NOTA

16. Recall that the sampling distribution of the adjusted sample standardized skewness coefficient follows a normal distribution with a mean of $\mu_{g_1} = E(g_1) = 0$ and a variance of $\sigma^2_{g_1} = Var(g_1) = \frac{6n(n-1)}{(n-2)(n+1)(n+3)}$ when sampling from a normally distributed population. One rule of thumb for rejecting the assumption that the population from which a sample is drawn is normal is if the absolute value of the adjusted sample standardized skewness coefficient exceeds 1. What is the approximate probability that we fail to reject the assumption that a SRS of size n = 20 comes from a normally distributed population based upon this rule of thumb? That is, what is the approximate probability that the absolute value of the adjusted sample standardized skewness coefficient does not exceed 1 when computed from a sample of size n = 20. Round your final answer to the thousandths place.

A: 0.883 B: 0.905 C: 0.949 D: 0.975 E: NOTA

17. The test that Señor Nieve gave the 12 students whose scores are in the table on the first page consisted of K multiple choice questions with 5 choices per question. The test scores in the table are on a standard 100-point scale where the lowest possible score is 0 and the highest possible score is 100; thus, the scores essentially represent the percentage of questions answered correctly out of the K questions in the test. Look closely at the test scores in the table and determine the number of questions on the test, K, given that each question is worth 4 "points" out of 100 total possible points. Then, calculate the probability that a student who randomly and independently guesses on each and every question on the test earns a score that is at least two standard deviations above the expected value for randomly guessing on the test. Round your final answer to four decimal places.

A: 0.0173 B: 0.0468 C: 0.9532 D: 0.9826 E: NOTA

18. Suppose Señor Nieve incorrectly scored the student who had the lowest score on the test and the correct score is a 72 instead of a 68. Which of the following three graphs from page 2 would change slightly in appearance if they were all still drawn with exactly the same scale and intervals on both axes as in the original graphs on page 2?

A: Graph A B: Graph B C: Graph C D: All of the above E: NOTA

19. A negative adjusted standardized skewness coefficient is evidence of skewness to the left, a positive one is evidence of skewness to the right, and if it is equal to 0, it is evidence of possible symmetry in the data; or equivalently, in the population from which the data came in each of these three cases. Additionally, given that at most one of the graphs on page 2 is possibly created from some other data set or information that is not contained in the table on page 1 (as noted in question #5), and that changing the lowest score will change the adjusted standardized skewness coefficient and may change the shape of at least one of the graphs on page 2; which of the following is the only absolutely correct classification of the original data set of 12 test scores on page 1 and why?

A: The distribution of the data is clearly perfectly symmetric since the mean is equal to the median.
B: The distribution of the data is clearly skewed to the right since Graph B is strongly skewed to the right.
C: The distribution of the data is slightly skewed to the right since Graph C is slightly skewed to the right.
D: The distribution of the data is clearly skewed to the left since the skewness coefficient is negative.
E: NOTA

20. Based upon your investigation and analysis of the data set of 12 test scores on page 1 throughout the previous questions, which of the following statements is / are false?

A: Whenever the mean of a data set is exactly equal to the median of the data set, a graph of the data set is always perfectly symmetric indicating that the distribution of the data set is perfectly symmetric.
B: Whenever the mean of a data set is greater than the median of the data set, a graph of the data set always appears skewed to the right indicating that the distribution of the data set is distinctly skewed to the right.
C: Whenever the mean of a data set is less than the median of the data set, a graph of the data set always appears skewed to the left indicating that the distribution of the data set is distinctly skewed to the left.
D: All of the above.
E: NOTA

21. Recall from question #11 that two students made up the test with scores of 60 and 84. Construct a 99% confidence interval for, μ, the true theoretical population mean of all possible test scores the entire class of 14 students could have earned on this test. You may assume all assumptions and conditions are met and that the population standard deviation is unknown. Round the final confidence interval limits to the thousandths place.

A: $77.659 < \mu < 90.341$ C: $75.158 < \mu < 92.842$ E: NOTA
B: $80.114 < \mu < 91.886$ D: $77.694 < \mu < 94.306$

22. Señor Nieve is quite pleased with his students' performance on this particular test this past year compared to how his students performed on the exact same test the previous year; however, he is curious to know if they performed statistically significantly better at the 5% level of significance. The previous year, his class of 20 students had a mean score of 74.2 with a standard deviation of 18.6480. Perform the appropriate statistical test to determine if his class of 14 students this past year performed significantly better than his class of 20 students the previous year on this same exact test. Calculate the sum of positive test statistic, the p-value, and the "exact" degrees of freedom of the test provided by your calculator program and add 1 to this sum if the results of the test support Señor Nieve's curiosity that his students performed significantly better this past year over the previous year and add 0 to the sum otherwise. Round all steps and the final answer to four decimal places (as needed), and again, you may assume that all inference assumptions and conditions for the requisite statistical test are met and that neither population standard deviation is known.

A: 34.2259 B: 33.2578 C: 33.2259 D: 34.2578 E: NOTA

23. Given that Señor Nieve is willing to view each of the two sample data sets of students' test scores in the previous question as a SRS from each respective theoretical population of all possible test scores, to whom can Señor Nieve validly generalize the results of the statistical test from the previous question?

A: All students who have ever taken AP Statistics throughout history.
B: All students who have ever taken that particular test throughout history.
C: Only those two particular sets of students who have taken any test during those two particular years.
D: Only those two particular sets of students and only on that particular test during those particular years.
E: NOTA

Use the following additional information to answer the next few questions:
Señor Nieve now wants to compare how the original set of 12 students whose test scores are in the table on page 1 performed on this particular test to their performance on the previous test in the class as well as perform a linear regression analysis of the set of paired test scores to see if there is a statistically significant linear relationship between the two sets of test scores. The test scores in the table on page 1 are for the Chapter 12 test and the previous test was on Chapter 11 (duh!). He is also eliminating the two students' test scores who made up the Chapter 12 test since they are a couple of clownboys who have the bad habit of always making up every test and so he does not want their scores to confound the results of his analyses. The scores on each test along with the sample means and the sample standard deviations for the 12 students are in the table below.

Student	A	B	C	D	E	F	G	H	I	J	K	L	Mean	SD
Ch. 11 Test	95	100	100	95	85	95	90	85	80	85	80	80	89.1667	7.6376
Ch. 12 Test	100	96	96	92	88	88	84	84	80	80	76	68	86	9.2638

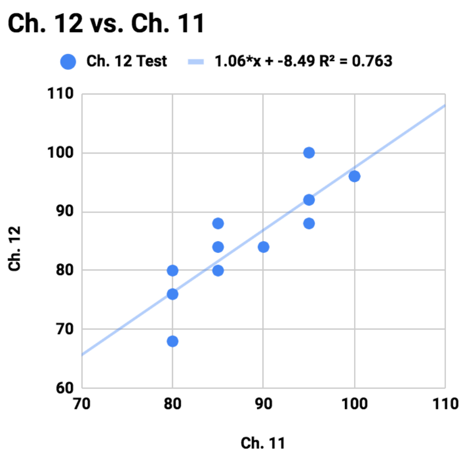
24. Perform the appropriate statistical test to determine if the mean difference in scores as determined by the Chapter 11 Test Scores minus the Chapter 12 Test Scores is significantly greater than 0 at the 5% level of significance. Compute the sum of the test statistic, the p-value, and the degrees of freedom of the test and then add 10 to this sum if the results of the test are statistically significant and add 5 otherwise. Round all steps and the final answer to four decimal places, as necessary, and you may assume that all inference assumptions and conditions are met.

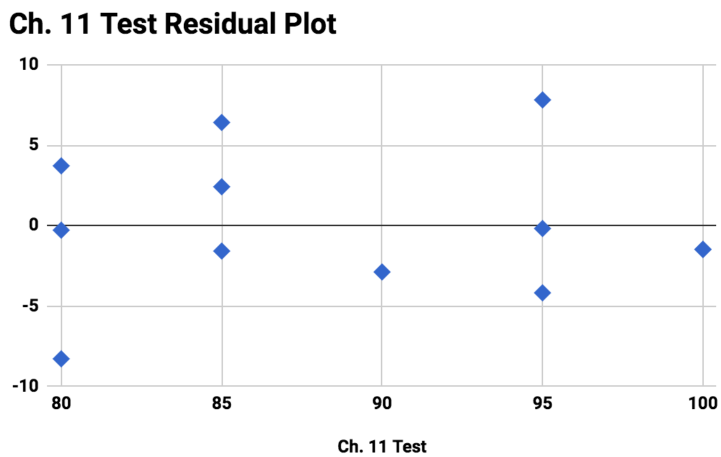
A: 23.4558 B: 27.3276 C: 23.4388 D: 27.5132 E: NOTA

Below are the *XLMiner* output tables displaying the results of the least squares linear regression analysis performed on the set of paired test scores on the previous page along with a scatterplot and a residual plot of the data. All output values are rounded to four decimal places for convenience.

Regression Statistics	
Pearson Correlation	0.8737
R Square	0.7634
Adjusted R Square	0.7397
Standard Error	4.7263
Observations	12

	Coefficients	*Standard Error*	*Test Statistic*	*Two-Tail P-value*
Intercept	-8.4935	16.6925	-0.5088	0.6219
Ch. 11 Test	1.0597	0.1866	5.6798	0.0002





25. Señor Nieve logically expects that there is a significant positive linear correlation between the pair of tests and that is clearly evident when we examine the results above. Using the results in the output table above, compute the sum of the test statistic, the p-value, and the degrees of freedom of the linear regression t-test of the hypotheses $H_0: \rho = 0 \; vs. H_A: \rho > 0$ where ρ (the Greek letter "rho") represents the population Pearson correlation coefficient between these two tests. Round all steps, parts, and the final answer to <u>four decimal places</u> and clearly, all inference assumptions and conditions are met.

A: 15.6800 B: 16.6799 C: 16.6800 D: 15.6799 E: NOTA

26. Recall from question #11 that two of Señor Nieve's clownboy students always make up every test. The student who scored a 60 on the Chapter 12 Test scored an 80 on the Chapter 11 Test and the student who scored an 84 on the chapter 12 Test scored a 100 (conveniently) on the Chapter 11 Test. Use the regression equation constructed from their 12 classmates scores in the table above to predict what each one would have earned on the Chapter 12 Test based upon each of their respective Chapter 11 Test scores according to this linear regression model then compute the sum of their residuals rounding <u>only the final answer</u> to the <u>nearest integer</u>.

A: –30 B: –44 C: 44 D: 30 E: NOTA

27. Which of the following statements are true based upon the linear regression analysis output tables above?

I. Approximately 76.34% of the variation in the Chapter 12 Test scores is explained by the linear relationship between the Chapter 12 Test scores and the Chapter 11 Test scores.
II. For every one-point increase in a student's Chapter 11 Test score, the estimated or predicted mean Chapter 12 Test score will increase by approximately 1.0597 points.
III. When using the linear regression model that is in the output tables above to predict a student's Chapter 12 Test score from his or her Chapter 11 Test score, on average, the observed Chapter 12 Test score will deviate from the predicted or estimated mean Chapter 12 Test score by approximately 4.7263 points.

A: I and II only B: I and III only C: II and III only D: All three E: NOTA

The following additional information is required for questions 28 and 29:
By this point in this test, I'm sure you are as sick of Señor Nieve as he is of you! However, he wishes to perform one final statistical analysis comparing his 14 students' performance on the AP Statistics Exam in this past year to how his 20 students performed on the AP Statistics Exam in the year prior. The table below displays the AP Statistics Exam scores for both sets of students. He wishes to know if there is convincing statistical evidence that the two score distributions are not distributed homogeneously at the 5% level of significance. (Note that even the two clownboys passed the exam!)

AP Statistics Exam Scores	3	4	5	Total
This Past Year	2	4	8	14
Prior Year	3	8	9	20
Total	5	12	17	34

28. If Señor Nieve is willing to view the two sets of scores as simple random samples from the respective theoretical population distributions of all possible scores the students could have earned, then all of the requisite assumptions and conditions for the required statistical test to determine if there is evidence that the two score distributions are not distributed homogeneously are met except possibly one. Which of the following is the reason for this possible violation of one of the conditions of the test?

A: The set of observed counts are not normally distributed.
B: The set of expected counts are not normally distributed.
C: Some of the observed counts are too small.
D: Some of the expected counts are too small.
E: NOTA

29. Perform the appropriate statistical test on the data in the table above in spite of the violation of the condition and determine if there is evidence that the AP Statistics Exam score distributions for the two sets of students during the two years are distributed heterogeneously. What is the sum of the test statistic, the p-value, and the degrees of freedom of the test rounded to <u>four decimal places</u> and then multiply the resulting sum by 2 if there is evidence of heterogeneity between the two AP Statistics Exam score distributions and add 2 to the sum otherwise.

A: 6.6198 B: 9.3099 C: 5.3099 D: 14.6198 E: NOTA

30. When Señor Nieve gets bored, he likes to perform a process where he takes data sets from past tests and treat them as if they were discrete random variables and then he takes very large simple random samples from them, with replacement. This process is very similar to a technique known as "bootstrapping." He then likes to take linear combinations of the results and calculate probabilities for even more fun. If he treats the two complete independent sets of scores on the Chapter 12 Test for each of the two groups of students we have been analyzing thus far as two "populations" (or equivalently two discrete random variables), then the expected value of performing this process on the set of test scores from this past year is 84 and the variance is 112. Likewise, when performing this process on the independent set of scores from the prior year of students, the expected value is 74.2 and the variance is 330.36. If Señor Nieve takes a SRS of 100 from each of these two independent "populations" (a.k.a. independent discrete random variables), what is the approximate probability that the means of the two samples are within 10 points of each other? Round only your final answer to <u>four decimal places</u>.

A: 0.3305 B: 0.5379 C: 0.6695 D: 0.4621 E: NOTA