Eastern West Virginia Community and Technical College
COURSE ASSESSMENT REPORT

Course Title and Number: MTH 225 Introduction to Statistics	**Academic Term and Year of Assessment Activity (Ex: Fall, 2014)** Fall 2015
Report Submitted By: Andrea Williams	**Number of Students Assessed:** 10
Date Report Submitted: 1/27/16	**Number of Sections Included:** 1
Course Delivery Format (list all modalities used in sections assessed. Ex: web based, VDL, traditional section, hybrid course, etc.): Traditional section	

Course Role in the Curriculum
Provide a description of the role the course serves in the curriculum (i.e. general education requirement, program technical core, restricted elective, etc.). Note all as appropriate.

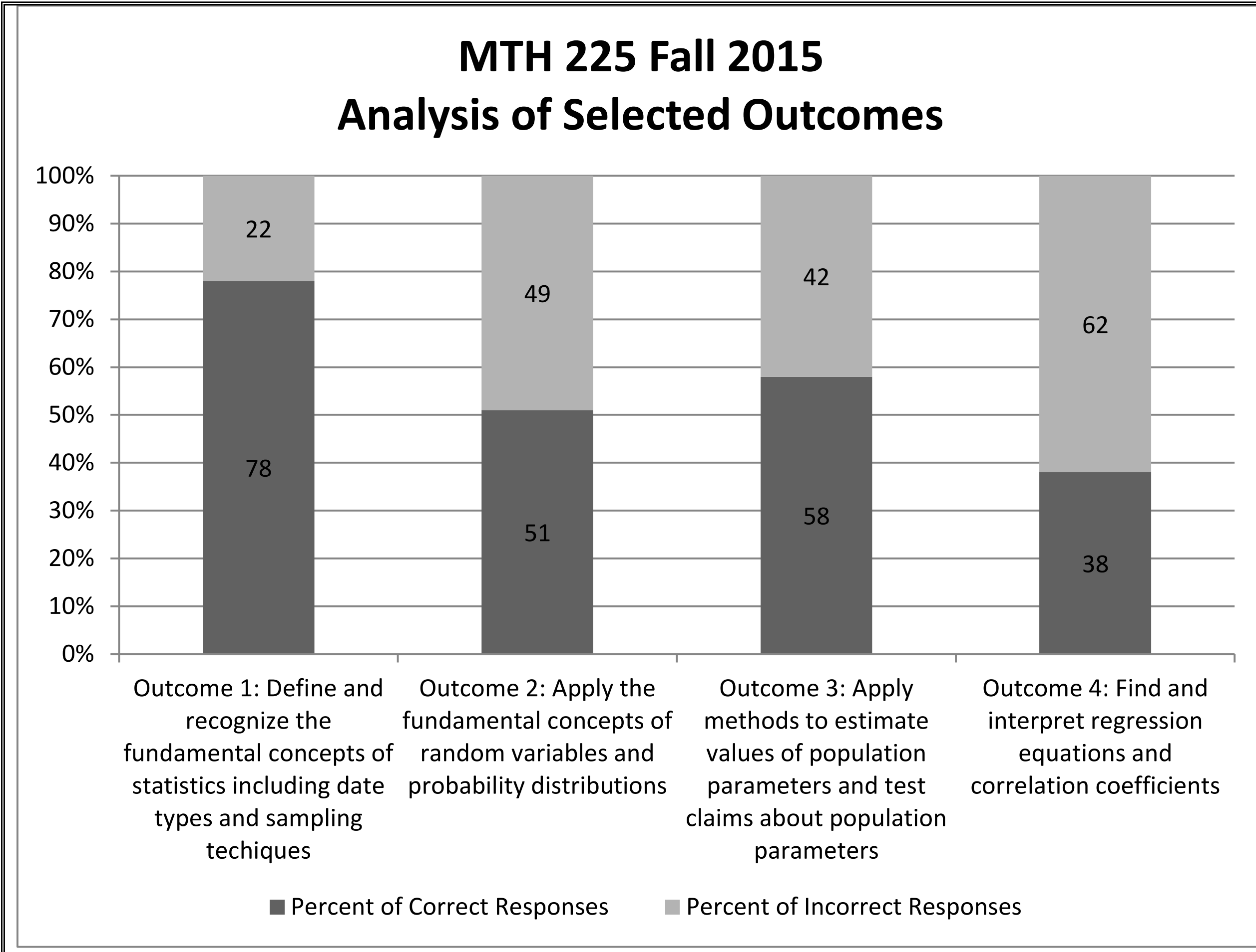
MTH 225 is provided to students as a transferable college-level math elective.

Assessment Methods
Provide a description of the assessment process used. Include description of instrument and performance standards in description. Note all methods.

Questions from unit tests throughout the semester are the basis for this assessment. Most of the questions were short answer questions, for which it was possible to receive partial credit, but for purposes of this analysis, only questions receiving full credit are considered correct with the exception of answers marked incorrect only because of a rounding error. Multiple questions are included in each outcome for analysis. A minimum satisfactory percent of correct responses for each outcome is 75%. Those failing to meet the standard are reviewed on an outcome-by-outcome basis.

Assessment Results
Provide a summary of results including tables/charts. Incorporate information from previous assessments as appropriate. Append additional pages if necessary. If appending, include notation in box to "See attached".

Four outcomes were analyzed, and only one met the 75% correct criterion. More details about the outcomes and the assessed questions are included in the action plan.



Course Level Assessment Summary of Outcomes, Indicators and Results
Course Title and Number: MTH 225 – Introduction to Statistics – Fall 2015
Number of students in assessment sample = 10
Number of Sections in Assessment = 1
Add additional rows to table if necessary

Learning Outcomes (Insert learning outcomes assessed during this cycle)	**Indicator (Insert indicators used for each outcome: exam question, scoring rubric, etc. Be specific)**	**Percent of Correct Responses +**	**Percent of Incorrect Responses**	**Performance Standard Met (75%) (yes or no)**
Outcome 1: Define and recognize the fundamental concepts of statistics including data types and sampling techniques	Test #1, Questions 1, 2*	78%	22%	Yes
Outcome 2: Apply the fundamental concepts of random variables and probability	Test #2, Questions 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22* Test #3, Questions 1, 2, 3, 4, 5, 6, 7, 8, 9, 10*	51%	49%	No

distributions				
Outcome 3: Apply methods to estimate values of population parameters and test claims about population parameters	Test #3, Questions 11, 13, 14, 15, 16, 17, 18, 19, 20* Test #4, Questions 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14* Test #5, Questions 7, 8, 9, 10*	58%	42%	No
Outcome 4: Find and interpret regression equations and correlation coefficients	Test #5, Questions 1, 2, 3, 4, 5, 6*	38%	62%	No

* See Attachment 1 for the content of each question.
+ See Attachment 1 for further analysis of performance by question. See Attachment 2 for further analysis of performance by sub-outcome.

Conclusions
Provide a brief summary of conclusions derived based on analysis of data. Append additional pages if necessary. If appending, include notation in box to "See attached".

Although the results are comparable to the previous course assessment, much can still be done for improvement. Changes since the previous assessment will continue to be implemented along with further emphasis on weaker topics as discussed below in the Action Plan.

Previous Assessment Reports and Results
Date of Previous Assessment: Spring 2013
List of Outcomes Not Met: Determine distribution probabilities, utilize the five-step hypothesis procedure, interpret correlation coefficient
Summary of Actions Taken to Address Unmet Learning Outcomes: Append additional pages if necessary. If appending, include notation in box to "See attached".

Note that the Master Course Record has been revised since the previous course assessment. Each of these outcomes now fall under a broader, course-level outcome:
- "determine distribution probabilities" is included in Outcome 2
- "utilize the five-step hypothesis procedure" is included in Outcome 3
- "interpret correlation coefficient" is included in Outcome 4.

When these outcomes are examined individually (see Attachment 2), the percent of correct responses becomes
- determine distribution probabilities: 40% (a 3% *decrease* from 2013)
- utilize the five-step hypothesis procedure: 76% (a 16% *increase* from 2013 and above the desired performance standard)
- interpret correlation coefficient: 70% (a 3% *decrease* from 2013)

When these individual results are considered, it shows that the recommendations from the 2013 course assessment have improved or at least maintained the previous level of performance. These recommendations will continue to be implemented along with the new recommendations found in the Action Plan.

The recommendations from the 2013 course assessment that were implemented were as follows:

- Review uniform distributions before the corresponding unit test.
- Emphasize the difference between finding a probability for a single piece of data and a probability for the mean of a set of data.
- Clarify the number of steps involved in a normal distribution problem to know when the solution is complete, and remind students of the key words to look for to know what type of problem it is.
- Remind students to double-check themselves when plugging information in to Statdisk.
- Create and distribute a handout with all the different rounding rules for the course.

Action Plan and Date for Reassessment

Identify action plan for improvement or maintaining current performance levels including outcomes identified for re-assessment, curriculum revision, LOT proposal, new or revised course activities to reinforce learning outcomes, etc. Append additional pages if necessary. If appending, include notation in box to "See attached".

Outcome 1: Define and recognize the fundamental concepts of statistics including data types and sampling techniques. This is actually a topic that is not discussed during class time since it is such a low-level objective; rather, the students are expected to read and learn this information from their textbook. A correct response rate of 78% confirms that this is a sufficient method for covering this particular material and will continue to be utilized in the future.

Outcome 2: Apply the fundamental concepts of random variables and probability distributions. This is one of the most important outcomes of the course but continues to be challenging for students. The questions that had the lowest number of correct responses included calculating probabilities involving a binomial distribution, finding a critical z value, recognizing the difference between a standard normal and a nonstandard normal distribution, finding probabilities involving normal distributions, and constructing a sampling distribution.

Constructing a sampling distribution is a very abstract topic; more consideration will be given as to whether this topic even needs to be covered. Will it create any confusion with subsequent sections if it is omitted?

Calculating the answer to a binomial distribution problem is quite simple *assuming* one recognizes that it is a binomial distribution problem, which few of the students did. The students did well on the one normal distribution problem that was not presented as a word problem. The questions that were put in the context of a real life application are the ones that caused the most trouble. The list of homework questions for these sections will be reviewed to ensure the students are encountering plenty of similar questions. The sample mean problem, which was an issue on the last course assessment, again received few correct responses because the students did not recognize that a different formula was required from the other normal distribution problems despite this being discussed at length in class.

Overall, to improve performance on this outcome, one possible solution is to create a review before each test with a group of problems *in random order* and discuss "How would you solve each of these problems?" In most cases it would not be necessary to completely solve the problem but just to get the students to recognize, "What type of problem is this? What formula do I need for this? What information do I need to extract from this problem to plug into Statdisk?" It is not so hard to answer these questions when one is working through the homework problems nicely categorized by section, but certainly more daunting when all the problems are grouped together on a test paper.

Outcome 3: Apply methods to estimate values of population parameters and test claims about population parameters. As discussed in the Previous Assessment Results, students have greatly improved in the area of testing claims. The questions that reduced the success rate of this outcome all involved finding critical values. This is clearly a topic for which more in-class examples and more homework problems are needed.

Outcome 4: Find and interpret regression equations and correlation coefficients. Two problem areas under this outcome included interpreting a coefficient of determination and determining the best predicted value. More in-class examples, homework problems, and review prior to the test are needed on these topics. The other two most commonly missed questions were about finding prediction intervals and non-linear regression equations. Both of these require a substantial amount of steps to arrive at the final solution. For future semesters, a homework problem similar to each will be graded to ensure the students are learning how to do this, but for the test, they will be given problems with some of the calculations already completed for them. Their responsibility will shift to interpreting the given information and using it to formulate the final solution.

Some other options that may improve overall performance in the course that will be considered before the course is offered again include

- Restructuring the course schedule. As is, every test includes two or three chapters. Less material on each test would likely improve performance. Unfortunately there are not enough class days in the semester to administer a test on every chapter, unless some take-home tests are given, but perhaps some of the more challenging chapters (in particular, Chapter 6 on normal distributions) could be taught as their own unit.
- Administering open-book, open-note tests. The students are provided so much information on each test already – a formula sheet, the handout on rounding rules, access to Statdisk, etc. – so it would not be much of a concession on the part of the instructor to allow open-book, open-note tests, but would likely reduce anxiety for the students thus improving performance.

Proposed date for reassessment is Fall 2017.

Assessment Committee Recommendation/Approval **(To be posted by Assessment Committee Chair)**
Ξ Approved as presented ☐ Approved with recommendations for future reports (Explanation Required) ☐ Resubmission Required. Reason for Resubmission: **Date: 3/23/16**

LOT Recommendation/Approval **(To be posted by Assessment Committee Chair)**
Ξ Approved as presented

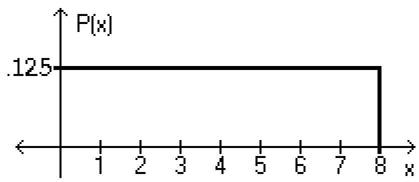
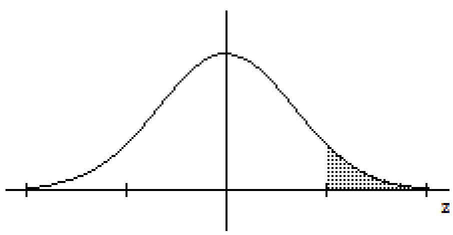
☐ Approved with recommendations for future reports (Explanation Required)
☐ Resubmission Required (Revision must be submitted to Assessment Committee before resubmitting to LOT). Reason for Resubmission:

Date: 4/18/16

Attachment 1: Performance by Question

<table>
<tr><th>Question</th><th>Percent of Correct Responses</th></tr>
<tr><td>1.1. Determine whether the given value is from a discrete or continuous data set: The number of stories in a Manhattan building is 22.
a. Discrete
b. Continuous</td><td>78%</td></tr>
<tr><td>1.2. Identify which type of sampling is used: A sample consists of every 49th student from a group of 496 students.
a. Stratified
b. Cluster
c. Convenience
d. Systematic
e. Random</td><td>78%</td></tr>
<tr><td>2.11. Determine whether the following is a probability distribution. If not, identify the requirement that is not satisfied.
<table><tr><td>x</td><td>P(x)</td></tr><tr><td>0</td><td>0.109</td></tr><tr><td>1</td><td>0.208</td></tr><tr><td>2</td><td>0.246</td></tr><tr><td>3</td><td>0.159</td></tr><tr><td>4</td><td>0.096</td></tr><tr><td>5</td><td>0.228</td></tr></table></td><td>80%</td></tr>
<tr><td>2.12. Identify the given random variable as being discrete or continuous: The number of freshmen in the required course, English 101.
a. Discrete
b. Continuous</td><td>90%</td></tr>
<tr><td>2.13. Find the mean of the given probability distribution: The random variable x is the number of houses sold by a realtor in a single month at the Sendsom’s Real Estate office. Its probability distribution is as follows.
<table><tr><td>Houses Sold (x)</td><td>Probability P(x)</td></tr><tr><td>0</td><td>0.24</td></tr><tr><td>1</td><td>0.01</td></tr><tr><td>2</td><td>0.12</td></tr><tr><td>3</td><td>0.16</td></tr><tr><td>4</td><td>0.01</td></tr><tr><td>5</td><td>0.14</td></tr><tr><td>6</td><td>0.11</td></tr><tr><td>7</td><td>0.21</td></tr></table></td><td>100%</td></tr>
<tr><td>2.14. Find the standard deviation for the given probability distribution.
<table><tr><td>x</td><td>P(x)</td></tr></table></td><td>60%</td></tr>
</table>

<table><tr><td>0</td><td>0.37</td></tr><tr><td>1</td><td>0.13</td></tr><tr><td>2</td><td>0.06</td></tr><tr><td>3</td><td>0.15</td></tr><tr><td>4</td><td>0.29</td></tr></table>	
2.15. Assume that a researcher randomly selects 14 newborn babies and counts the number of girls selected, x. The probabilities corresponding to the 14 possible values of x are summarized in the given table. Answer the question using the table: **Probabilities of Girls** <table><tr><th>x(girls)</th><th>P(x)</th><th>x(girls)</th><th>P(x)</th><th>x(girls)</th><th>P(x)</th></tr><tr><td>0</td><td>0.000</td><td>5</td><td>0.122</td><td>10</td><td>0.061</td></tr><tr><td>1</td><td>0.001</td><td>6</td><td>0.183</td><td>11</td><td>0.022</td></tr><tr><td>2</td><td>0.006</td><td>7</td><td>0.209</td><td>12</td><td>0.006</td></tr><tr><td>3</td><td>0.022</td><td>8</td><td>0.183</td><td>13</td><td>0.001</td></tr><tr><td>4</td><td>0.061</td><td>9</td><td>0.122</td><td>14</td><td>0.000</td></tr></table> Find the probability of selecting 9 or more girls.	90%
2.16. Assume that there is a 0.15 probability that a basketball playoff series will last four games, a 0.30 probability that it will last five games, a 0.25 probability that it will last six games, and a 0.30 probability that it will last seven games. Is 7 an usually high number of games for a series? Justify your answer.	30%
2.17. Determine whether the given procedure results in a binomial distribution. If not, state the reason why. Rolling a single die 59 times, keeping track of the numbers that are rolled. a. Not binomial: there are too many trials. b. Not binomial: there are more than two outcomes for each trial. c. Not binomial: the trials are not independent. d. Procedure results in a binomial distribution.	100%
2.18. Find the probability of at least 2 girls in 5 births. Assume that male and female births are equally likely and that the births are independent events.	0%
2.19. In a survey of 300 college graduates, 56% reported that they entered a profession closely related to their college major. If 8 of those survey subjects are randomly selected without replacement for a follow-up survey, what is the probability that 3 of them entered a profession closely related to their college major?	10%
2.20. According to a college survey, 22% of all students work full time. Find the mean for the number of students who work full time in samples of size 16.	100%
2.21. On a multiple choice test with 17 questions, each question has four possible answers, one of which is correct. For students who guess at all answers, find the standard deviation for the number of correct answers.	50%
2.22. Determine if the outcome is unusual. Consider as unusual any result that differs from the mean by more than 2 standard deviations. That is, unusual values are either less	50%

than $\mu - 2\sigma$ or greater than $\mu + 2\sigma$. A survey for brand recognition is done and it is determined that 68% of consumers have heard of Dull Computer Company. A survey of 800 randomly selected consumers is to be conducted. For such groups of 800, would it be unusual to get 634 consumers who recognize the Dull Computer Company name? Justify your answer.	
3.1. Using the following uniform density curve, answer the question.  What is the probability that the random variable has a value less than 7?	50%
3.2. If z is a standard normal variable, find the probability that z lies between -1.10 and -0.36.	50%
3.3. Find the indicated z score. The graph depicts the standard normal distribution with mean 0 and standard deviation 1. Shaded area is 0.0694. 	80%
3.4. Find the indicated value: $z_{0.10}$.	20%
3.5. List two characteristics that distinguish a standard normal distribution from a nonstandard normal distribution.	30%
3.6. Assume that X has a normal distribution, and find the indicated probability. The mean is $\mu = 60.0$ and the standard deviation is $\sigma = 4.0$. Find the probability that X is less than 53.0.	60%
3.7. The weekly salaries of teachers in one state are normally distributed with a mean of \$490 and a standard deviation of \$45. What is the probability that a randomly selected teacher earns more than \$525 a week?	20%
3.8. Scores on a test are normally distributed with a mean of 63.2 and a standard deviation of 11.7. Find P_{81}, which separates the bottom 81% from the top 19%.	20%
3.9. Three randomly selected households are surveyed as a pilot project for a larger survey to be conducted later. The numbers of people in the households are 1, 4, and 7. Consider the values of 1, 4, and 7 to be a population. Assume that samples of size n = 2 are randomly selected with replacement from the population of 1, 4 Three randomly selected households are surveyed as a pilot project for a larger survey to be conducted later. The numbers of people in the households are 1, 4, and 7. Consider the values of 1,	10%

4, and 7 to be a population. Assume that samples of size n = 2 are randomly selected with replacement from the population of 1, 4, and 7. The nine different samples are as follows: (1,1), (1,4), (1,7), (4,1), (4,4), (4,7), (7,1), (7,4), and (7,7). (i) Construct a probability distribution table that describes the sampling distribution of the proportion of even numbers when samples of size n = 2 are randomly selected. (ii) Does the mean of the sample proportions target the value of the population proportion? (iii) Does the sample proportion make a good estimator of the population proportion?	
3.10. The amount of snowfall falling in a certain mountain range is normally distributed with a mean of 70 inches, and a standard deviation of 10 inches. What is the probability that the mean annual snowfall during 25 randomly picked years will exceed 72.8 inches?	20%
3.11. Find the critical value $z_{\alpha/2}$ that corresponds to a 96% confidence level.	10%
3.13. Use the given degree of confidence and sample data to construct a confidence interval for the population proportion p. Round endpoints to three significant figures. Then write a sentence explaining what this confidence interval means. When 268 college students are randomly selected and surveyed, it is found that 114 own a car. Find a 99% confidence interval for the true proportion of all college students who own a car.	60%
3.14. Use the given data to find the minimum sample size required to estimate the population proportion. Margin of error: 0.04; confidence level: 99%; from a prior study, $\hat{p}$ is estimated by 0.12.	80%
3.15. Do one of the following, as appropriate: (a) Find the critical value $z_{\alpha/2}$, (b) find the critical value $t_{\alpha/2}$, (c) state that neither the normal nor the t distribution applies. 99%; n = 17; σ is unknown; population appears to be normally distributed.	10%
3.16. Use the confidence level and sample data to find a confidence interval for estimating the population μ. Round endpoints to the same number of decimal places as the sample mean. Then write a sentence explaining what this confidence interval means. A random sample of 73 light bulbs had a mean life of $\bar{x} = 413$ hours with a standard deviation of $s = 33$ hours. Construct a 90% confidence interval for the mean life, μ, of all light bulbs of this type.	50%
3.17. Use the given degree of confidence and sample data to construct a confidence interval for the population mean μ. Assume that the population has a normal distribution. Round endpoints to one more place than the given data values. Then write a sentence explaining what this confidence interval means. The football coach randomly selected ten players and times how long each player took to perform a certain drill. The times (in minutes) were: 7.9 10.1 9.8 8.5 11.7 7.4 6.0 11.5 10.3 12.8 Determine a 95% confidence interval for the mean time for all players.	40%
3.18. Use the given information to find the minimum sample size required to estimate an unknown population mean μ. How many commuters must be randomly selected to estimate the mean driving time of Chicago commuters? We want 98% confidence that the sample mean is within 2 minutes of the population mean, and the population standard	60%

deviation is known to be 10 minutes.	
3.19. Find the critical value χ_R^2 corresponding to a sample size of 12 and a confidence level of 98 percent.	30%
3.20. Use the given degree of confidence and sample data to find a confidence interval for the population standard deviation σ. Assume that the population has a normal distribution. Round the confidence interval limits to the same number of decimal places as the sample standard deviation. Then write a sentence explaining what this confidence interval means. College students' annual earnings: 98% confidence; n = 9, $\bar{x} = \$3361$, s = \$865.	40%
4.1. Express the null hypothesis and the alternative hypothesis in symbolic form. Use the correct symbol (μ, p, σ) for the indicated parameter. Then determine whether the test would be left-tailed, right-tailed, or two-tailed. A psychologist claims that more than 5.8 percent of the population suffers from professional problems due to extreme shyness. Use p, the true percentage of the population that suffers from extreme shyness.	44%
4.2. A medical researcher claims that 10% of children suffer from a certain disorder. Identify the type I error for the test. a. Reject the claim that the percentage of children who suffer from the disorder is different from 10% when that percentage really is different from 10%. b. Fail to reject the claim that the percentage of children who suffer from the disorder is equal to 10% when that percentage is actually different from 10%. c. Fail to reject the claim that the percentage of children who suffer from the disorder is equal to 10% when that percentage is actually 10%. d. Reject the claim that the percentage of children who suffer from the disorder is equal to 10% when that percentage is actually 10%.	78%
4.3. Identify which distribution is used for the given claim. Then find the value of the test statistic using the appropriate formula. A claim is made that the proportion of children who play sports is less than 0.5, and the sample statistics include n = 1320 subjects with 30% saying that they play a sport.	22%
4.4. Use the given information to find the P-value. Then us a 0.05 significance level and state the conclusion about the null hypothesis (reject the null hypothesis or fail to reject the null hypothesis). The test statistic in a two-tailed test is z = -1.63.	78%
4.5. Assume that the data has a normal distribution. Find the critical z value used to test a null hypothesis. Then assume the claim produces a test statistic of z = 1.25 and state the conclusion about the null hypothesis (reject the null hypothesis or fail to reject the null hypothesis). $\alpha = 0.09$ for a right-tailed test.	22%
4.6. Identify the null hypothesis, alternative hypothesis, test statistic, P-value, conclusion about the null hypothesis, and final conclusion that addresses the original claim. In a sample of 167 children selected randomly from one town, it is found that 37 of them suffer from asthma. At the 0.05 significance level, test the claim that the proportion of all children in the town who suffer from asthma is 11%.	89%

4.7. Identify the null hypothesis, alternative hypothesis, test statistic, P-value, conclusion about the null hypothesis, and final conclusion that addresses the original claim. A manufacturer makes ball bearings that are supposed to have a mean weight of 30 g. A retailer suspects that the mean weight is actually less than 30 g. The mean weight for a random sample of 16 ball bearings is 28.6 g with a standard deviation of 4.4 g. At the 0.05 significance level, test the claim that the sample comes from a population with a mean weight less than 30 g.	100%
4.8. Identify the null hypothesis, alternative hypothesis, test statistic, P-value, conclusion about the null hypothesis, and final conclusion that addresses the original claim. A manufacturer uses a new production method to produce steel rods. A random sample of 17 steel rods resulted in lengths with a standard deviation of 4.7 cm. At the 0.10 significance level, test the claim that the new production method has lengths with a standard deviation different from 3.5 cm, which was the standard deviation for the old method.	67%
4.9. Identify the null hypothesis, alternative hypothesis, test statistic, P-value, conclusion about the null hypothesis, and final conclusion that addresses the original claim. In a random sample of 500 people aged 20-24, 22% were smokers. In a random sample of 450 people aged 25-29, 14% were smokers. Test the claim that the proportion of smokers in the two age groups is the same. Use a significance level of 0.01.	67%
4.10. Construct the indicated confidence interval for the difference between population proportions $p_1 - p_2$. Assume that the samples are independent and that they have been randomly selected. Give the endpoints of the interval to three significant figures. In a random sample of 300 women, 45% favored stricter gun control legislation. In a random sample of 200 men, 25% favored stricter gun control legislation. Construct a 98% confidence intereval for the difference between the population proportions $p_1 - p_2$.	78%
4.11. Identify the null hypothesis, alternative hypothesis, test statistic, P-value, conclusion about the null hypothesis, and final conclusion that addresses the original claim. A researcher wishes to determine whether people with high blood pressure can reduce their blood pressure, measured in mm Hg, by following a particular diet. Use a significance level of 0.01 to test the claim that the treatment group is from a population with a smaller mean than the control group. Treatment Group: $n_1 = 35$, $\bar{x}_1 = 189.1$, $s_1 = 38.7$ Control Group: $n_2 = 28$, $\bar{x}_2 = 203.7$, $s_2 = 39.2$	67%
4.12. A paint manufacturer made a modification to a paint to speed up its drying time. Independent simple random samples of 11 cans of type A (the original paint) and 9 cans of type B (the modified paint) were selected and applied to similar surfaces. The drying times, in hours, were recorded. The summary statistics are as follows. Type A: $n_1 = 11$, $\bar{x}_1 = 76.3$ hrs, $s_1 = 4.5$ hrs Type B: $n_2 = 9$, $\bar{x}_2 = 65.1$ hrs, $s_2 = 5.1$ hrs The following 98% confidence interval was obtained for $\mu_1 - \mu_2$, the difference between the mean drying time for paint cans of type A and the mean drying time for paint cans of	67%

type B: 4.90 hrs $< \mu_1 - \mu_2 <$ 17.50 hrs. What does the confidence interval suggest about the population means?
a. The confidence interval includes only positive values which suggests that the two population means might be equal. There doesn't appear to be a significant difference between the mean drying time for paint type A and the mean drying time for pain type B. The modification does not seem to be effective in reducing drying times.
b. The confidence interval includes 0 which suggests that the two population means might be equal. There doesn't appear to be a significant difference between the mean drying time for paint type A and the mean drying time for paint type B. The modification does not seem to be effective in reducing drying times.
c. The confidence interval includes only positive values which suggests that the mean drying time for pain type A is smaller than the mean drying time for paint type B. The modification does not seem to be effective in reducing drying times.
d. The confidence interval includes only positive values which suggests that the mean drying time for paint type A is greater than the mean drying time for pain type B. The modification seems to be effective in reducing drying times.

4.13. Construct a confidence interval for μ_d, the mean of the differences d for the population of paired data. Assume that the population of paired differences is normally distributed. A coach uses a new technique in training middle distance runners. The times for 9 different athletes to run 800 meters before and after this training are shown below.

89%

Athlete	A	B	C	D	E	F	G	H	I
Time before training (seconds)	115.2	120.9	108.0	112.4	107.5	119.1	121.3	110.8	122.3
Time after training (seconds)	116.0	119.1	105.1	111.9	109.1	115.2	118.5	110.7	120.9

Construct a 99% confidence interval for the mean difference of the "before" minus "after" times. Round the endpoints to one more decimal place than the given data values.

4.14. Identify the null hypothesis, alternative hypothesis, test statistic, P-value, conclusion about the null hypothesis, and final conclusion that addresses the original claim. When 25 randomly selected customers enter any one of several waiting lines, their waiting times have a standard deviation of 5.8 minutes. When 16 randomly selected customers enter a single main waiting line, their waiting times have a standard deviation of 2.19 minutes. Use a 0.05 significance level to test the claim that there is more variation in the waiting times when several lines are used.

67%

5.7. Given below are the analysis of variance results from a Minitab display. Assume that you want to use a 0.05 significance level in testing the null hypothesis that the different samples come from populations with the same mean. What can you conclude about the equality of the population means?

80%

Source	DF	SS	MS	F	p
Factor	3	30	10.00	1.6	0.264
Error	8	50	6.25		
Total	11	80			

a. Reject the null hypothesis since the P-value is greater than the significance level.
b. Accept the null hypothesis since the P-value is less than the significance level.
c. Reject the null hypothesis since the P-value is less than the significance level.
d. Accept the null hypothesis since the P-value is greater than the significance level.

5.8. Test the claim that the samples come from populations with the same mean. Assume that the populations are normally distributed with the same variance. Random samples of

70%

four different models of cars were selected and the gas mileage of each car was measured. The results are shown below. Model A / Model B / Model C / Model D 23 / 28 / 30 / 25 25 / 26 / 28 / 26 24 / 29 / 32 / 25 26 / 30 / 27 / 28 Test the claim that the four different models have the same population mean. Use a significance level of 0.05.	
5.9. A manager records the production output of three employees who each work on three different machines for three different days. The sample results are given below and the Minitab results follow.	60%

Employee

		A	B	C
	I	23, 27, 29	30, 27, 25	18, 20, 22
Machine	II	25, 26, 24	24, 29, 26	19, 16, 14
	III	28, 25, 26	25, 27, 23	15, 11, 17

ANALYSIS OF VARIANCE ITEMS

SOURCE	DF	SS	MS	F	P
MACHINE	2	34.67	17.33	3.1837	0.066
EMPLOYEE	2	504.67	252.33	46.3469	0.000
INTERACTION	4	26.67	6.67	1.22	0.335
ERROR	18	98.00	5.44		
TOTAL	26	664.00			

Using a 0.05 significance level, apply the methods of two-way analysis of variance. What do you conclude?

5.10. The following data show annual income, in thousands of dollars, categorized according to the two factors of gender and level of education. Using a 0.05 significance level, apply the methods of two-way analysis of variance. What do you conclude? Round P-values to four decimal places. — 60%

	Female	Male
High school	23, 27, 24, 26	25, 26, 22, 24
College	28, 36, 31, 33	35, 32, 39, 28
Advanced degree	41, 38, 43, 49	35, 50, 47, 44

5.1. Find the value of the linear correlation coefficient r. Then find the P-value, and, using a 0.05 significance level, determine whether there is sufficient evidence to support a claim of a linear correlation between the two variables. The paired data below consist of the test scores of 6 randomly selected students and the number of hours they studied for the test. — 70%

Hours	5	10	4	6	10	9
Score	64	86	69	86	59	87

5.2. Use the given data to find the equation of the regression line. Round the final values to three significant digits, if necessary. — 70%

x	24	26	28	30	32
y	15	13	20	16	24

5.3. Use the given data to find the coefficient of determination and the percentage of the total variation that can be explained by the linear relationship between the two variables. Write a sentence explaining what that percentage means. The paired data below consists of test scores and hours of preparation for 5 randomly selected students.

30%

x Hours of preparation	5	2	9	6	10
y Test score	64	48	72	73	80

5.4. The equation of the regression line for the paired data below is $\hat{y} = 6.1829 + 4.3394x$ and the standard error of estimate is $s_3 = 1.6419$. Find the 99% prediction interval of y for x = 12. Round the sample mean to one more decimal place than the given data values; round the endpoints of the interval to the same number of decimal places as the given data values.

30%

x	9	7	2	3	4	22	17
y	43	35	16	21	23	102	81

5.5. Based on the data from six students, the regression equation relating number of hours of preparation (x) and test score (y) is $\hat{y} = 67.3 + 1.07x$. The same data yield r = 0.224 and $\bar{y} = 75.2$. What is the best predicted test score for a student who spent 2 hours preparing for the test?

10%

5.6. Identify the mathematical model that best fits the data. Assume that the model is to be used only for the scope of the given data and consider only linear, quadratic, logarithmic, exponential, and power models. Use Desmos to compare values of R^2 and obtain the regression equation of the model that best fits the data.

20%

x	1	2	3	4	5
y	3	9	17	30	40

Attachment 2: Performance by Sub-Outcome

Outcome	Sub-Outcome	Test Question(s)	Percent of Correct Responses
Outcome 1: Define and recognize the fundamental concepts of statistics including data types and sampling techniques	Determine whether a value is from a discrete or continuous data set	1.1	78%
	Classify sampling methods	1.2	78%
Outcome 2: Apply the fundamental concepts of random variables and probability distributions	Recognize whether a table gives a probability distribution	2.11	80%
	Identify whether a given value is a discrete random variable, continuous random variable, or not a random variable	2.12	90%
	Calculate parameters of a probability distribution	2.13, 2.14	80%
	Calculate probabilities using a probability distribution	2.15	90%
	Apply the range rule of thumb to distinguish between usual and unusual values of a random variable	2.22	50%
	Use probabilities to determine whether results are usual	2.16	30%
	Recognize whether a probability distribution is a binomial distribution	2.17	100%
	Calculate probabilities using a binomial distribution	2.18, 2.19	5%
	Calculate parameters of a binomial distribution	2.20, 2.21	75%
	Use a density curve to determine probabilities	3.1	50%
	Distinguish between a standard and a nonstandard normal distribution	3.5	30%
	Find the area or probability corresponding to a given	3.2, 3.6, 3.7	43%

	z-score		
	Find the z-score corresponding to a given area or probability	3.3, 3.8	50%
	Find critical z values	3.4	20%
	Create and interpret sampling distributions	3.9	10%
	Apply the central limit theorem	3.10	20%
Outcome 3: Apply methods to estimate values of population parameters and test claims about population parameters	Find critical z, t, and chi-square values	3.11, 3.15, 3.19	17%
	Construct and interpret confidence intervals	3.13, 3.16, 3.17, 3.20	48%
	Compute the sample size needed to estimate a parameter	3.14, 3.18	70%
	Identify the null hypothesis and alternative hypothesis from a given claim, and express both in symbolic form	4.1, 4.6, 4.7, 4.8	75%
	Identify Type I and Type II errors that correspond to a given claim	4.2	78%
	Identify the appropriate distribution for a given claim and find the value of the test statistic	4.3	22%
	Identify what type of test applies to a given hypothesis	4.1	44%
	Find the P-value and critical value(s) for a given claim	4.4, 4.5, 4.6, 4.7, 4.8	71%
	State a conclusion based on a P-value or critical value using the appropriate wording	4.4, 4.5, 4.6, 4.7, 4.8	71%
	Construct and interpret confidence intervals comparing parameters of two populations	4.10, 4.12, 4.13	78%
	Test a claim involving parameters of two populations	4.9, 4.11, 4.14	67%
	Use one-way analysis of variance (ANOVA) to test a claim about the	5.7, 5.8	75%

	equality of means of three or more populations		
	Use two-way ANOVA to test for an interaction between two factors and to determine whether each factor has an effect on a set of data	5.9, 5.10	60%
Outcome 4: Find and interpret regression equations and correlation coefficients	Calculate and interpret a correlation coefficient	5.1	70%
	Find a regression equation	5.2	70%
	Determine the best predicted value	5.5	10%
	Calculate and interpret a prediction interval	5.4	30%
	Determine which nonlinear regression equation best models a set of data	5.6	20%