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 2017
Report Submitted By: Andrea Williams	**Number of Students Assessed:** 7
Date Report Submitted: 1/4/18	**Number of Sections Included:** 1
Course Delivery Format (list all modalities used in sections assessed. Ex: web based, VDL, traditional section, hybrid course, etc.): Web-based	

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.

Homework assignments and unit tests are used as the basis for this assessment. All homework assignments were given through MyStatLab, an online homework and testing platform, where students had access to numerous learning aids and had unlimited attempts at each question. The instructor provided video lectures in Blackboard of each section that students were encouraged to watch before attempting the homework. Although many questions had multiple parts, for purposes of this analysis, only questions receiving full credit are considered correct.

The unit tests were given on paper and proctored on campus. Students were allowed to use textbooks, notes, a calculator, and, for the last three tests, the statistical analysis package Statdisk. 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.

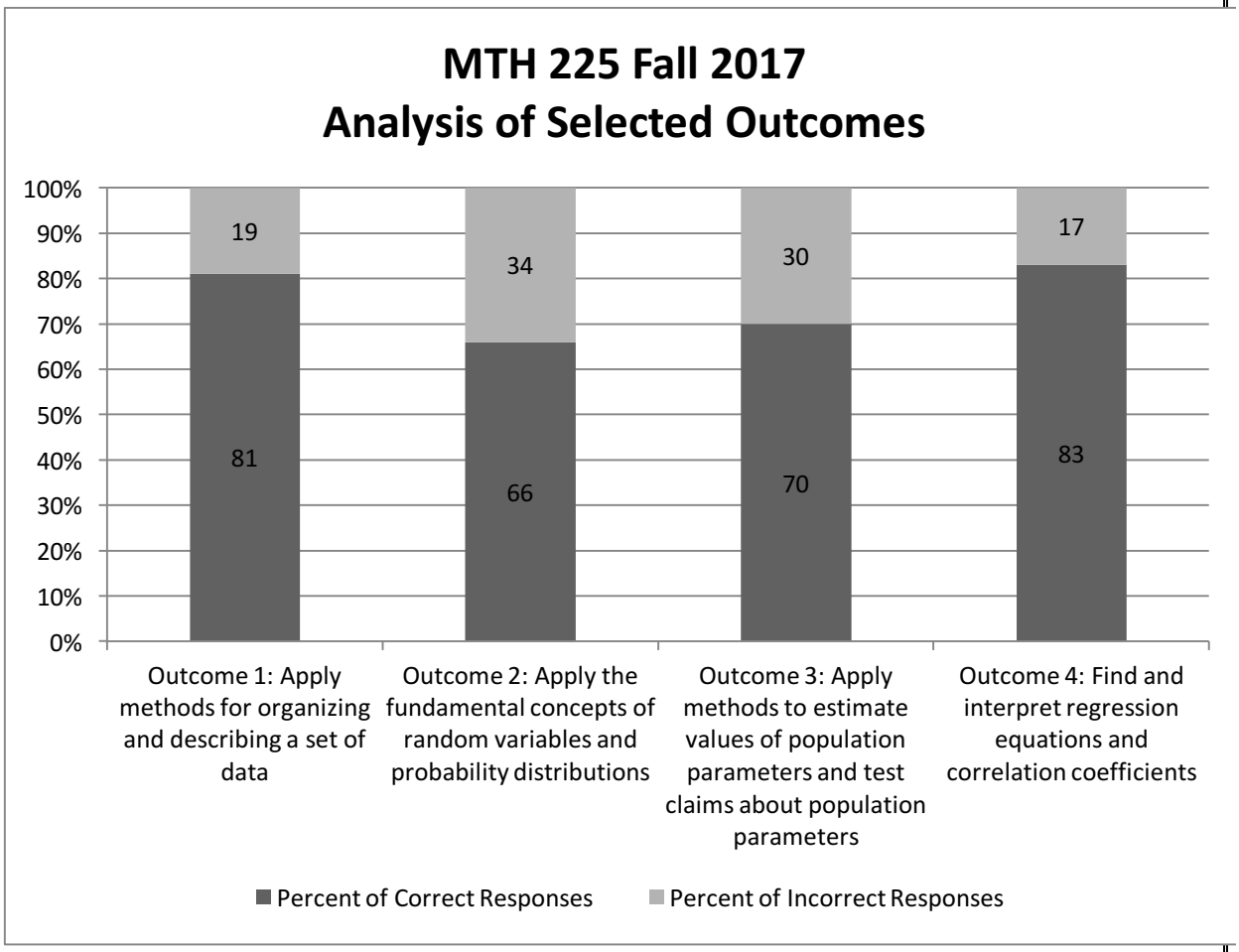
Multiple questions are included in each outcome for analysis. A minimum satisfactory percent of correct responses for each outcome is 75%.

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, the three that failed to meet the minimum performance standard on the last assessment and one new outcome that was not assessed previously. Two outcomes out of the four met the 75% correct criterion; the two that failed to meet the standard showed substantial improvement over the last assessment. However, this was the first time an online section of this course was

assessed, so results may not be comparable with previous cohorts.



Course Level Assessment Summary of Outcomes, Indicators and Results 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: Apply methods for organizing and describing a set of data	Homework Sections 2.1, 2.2, 2.3, 2.4, 3.1, 3.2, 3.3 Test Questions 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 1.10, 1.11, 1.12, 1.13, 1.14, 1.15, 1.16, 1.17, 1.18, 1.19, 1.20, 1.21, 1.22, 1.23, 1.24, 1.25	81%	19%	Yes
Outcome 2: Apply the fundamental	Homework Sections 5.1, 5.2, 6.1, 6.2, 6.3, 6.4	66%	34%	No

concepts of random variables and probability distributions	Test Questions 2.8, 2.9, 2.10, 2.11, 2.12, 2.13, 2.14, 2.15, 2.16, 2.17, 2.18, 2.19, 2.20			
Outcome 3: Apply methods to estimate values of population parameters and test claims about population parameters	Homework Sections 7.1, 7.2, 7.3, 8.1, 8.2, 8.3, 8.4, 9.1, 9.2, 9.3, 9.4, 12.1, 12.2 Test Questions 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, 3.9, 3.10, 3.11, 3.12, 3.13, 3.14, 3.15, 3.16, 3.17, 3.18, 3.19, 3.20, 4.8, 4.9	70%	30%	No
Outcome 4: Find and interpret regression equations and correlation coefficients	Homework Sections 10.1, 10.2, 10.3, 10.4, 10.5 Test Questions 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7	83%	17%	Yes

* Please note if using a different minimum performance standard.
+ See Attachment 1 for questions and outcomes.
^ Homework assignments were worth 25% of the final course grade; unit tests were worth 35%. The percent of correct responses was calculated as a weighted average based on these percentages. See Attachment 2 for results separated by performance indicator.

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".

As mentioned above, results from this online section may not be comparable with previous face-to-face cohorts. However, it does appear that revisions made since the last assessment (see below) have been effective, so they will continue to be implemented along with the new suggestions in the action plan below.

This was also the first time that homework assignments have been included in the course assessment in addition to the unit tests. Incorporating more performance indicators provides a better picture of what the students gained from the course. Attachment 2 shows that students did significantly better on the homework assignments than on the unit tests, so efforts for improving student performance need to focus on the tests. This is also addressed in the action plan below.

Previous Assessment Reports and Results

Date of Previous Assessment: Fall 2015

List of Outcomes Not Met:

- Apply the fundamental concepts of random variables and probability distributions
- Apply methods to estimate values of population parameters and test claims about population parameters
- Find and interpret regression equations and correlation coefficients

Summary of Actions Taken to Address Unmet Learning Outcomes: Append additional pages if necessary. If appending, include notation in box to "See attached".

While two of the outcomes still were not met at the desired 75% performance standard, all three showed substantial improvement over the Fall 2015 assessment with increases ranging from 12% to 45%. The outcome that saw the greatest improvement and is now above 75% was "find and interpret regression equations and correlation coefficients." This increase is likely due to the suggestion from the previous assessment to avoid laborious calculations on the test and focus on interpretation.

Also as suggested in the previous assessment, the list of homework problems for the course was reviewed; extra problems and examples were added for topics known to be areas of difficulty. For example, finding critical values, one specific area of concern on the last assessment, was emphasized over several chapters, and any time an example required finding a critical value, the steps were written out in detail.

It was decided that the abstract topic of sampling distributions *is* necessary for understanding subsequent topics, but students were assessed on sampling distributions exclusively by a homework assignment rather than on a test.

The students have been given open-book, open-note tests the last two times the course has been offered, and this has helped performance to a certain extent and will continue to be utilized. However, the list of topics that the students find challenging remains consistent across semesters, with or without books and notes.

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".

As Attachment 2 shows, students performed significantly better on homework assignments than unit tests, regardless of outcome. To improve performance on tests and in the class overall, a review session should be offered before each test. While the course schedule does not allow time for this in a live class, optional out-of-class reviews could be offered, at least until their level of effectiveness is determined. For online sections of the course, the class could be offered instead as a hybrid, utilizing the required class time for questions and review. A hybrid course would also give the students the opportunity to present their final project to their classmates, a cohesive element lacking from the strictly online section this semester.

This was the first semester MyStatLab was utilized in the course, and it proved to be an effective learning tool. It will continue to be used for at least part of the homework for the course, even in live sections. Consideration will also be given as to how MyStatLab can be integrated into the aforementioned test reviews.

Proposed date for reassessment is Fall 2017. The two outcomes not meeting the 75% performance standard will be reassessed along with two new outcomes.

Assessment Committee Recommendation/Approval

(To be posted by Assessment Committee Chair)
☒ Approved as presented ☐ Approved with recommendations for future reports (Explanation Required) ☐ Resubmission Required. Reason for Resubmission: **Date: 02/09/2018**

Attachment 1 – Performance Indicators Used for Assessment

Homework

In the interest of space (some outcomes included over 100 homework questions), individual homework questions will not be included in this report; instead, the objectives covered in each section as listed in MyStatLab are given below.

Section	Objectives Covered
2.1	• Understand the concepts related to distributions. • Summarize data in the format of a frequency distribution and a relative frequency distribution. • Identify values of class width, class midpoint, class limits, and class boundaries. • Determine whether a frequency distribution is approximately a normal distribution. • Summarize data in the format of a frequency distribution and a relative frequency distribution. • Construct a cumulative frequency distribution.
2.2	• Picture the distribution of data in the format of a histogram and examine.
2.3	• Understand the concepts related to graphing data. • Develop an ability to graph data using a dotplot. • Develop an ability to graph data using a stemplot. • Develop an ability to graph data using a time-series graph. • Develop an ability to graph data using a Pareto chart. • Develop an ability to graph data using a pie chart. • Develop an ability to graph data using a frequency polygon. • Determine when a graph is deceptive.
2.4	• Understand the concepts related to scatterplots and correlation. • Construct and analyze a scatterplot of paired data.
3.1	• Develop the ability to measure the center of data by finding the mean, median, mode, and midrange. • Determine whether an outlier has a substantial effect on the mean and median.
3.2	• Interpret values of the standard deviation by applying the range rule of thumb. • Find values of the range, variance, and standard deviation. • Understand the concepts related to measures of variation. • Find the coefficient of variation of two samples and then compare the results. • Use the empirical rule or Chebyshev\'s theorem to estimate the variation in a sample.
3.3	• Compute a z score and use it to determine whether a value is significant. • Develop the ability to construct a boxplot from a set of data. • Identify percentile values and quartile values from a set of data.
5.1	• Understand concepts related to random variables and probability distributions. • Determine when a potential probability distribution actually satisfies the necessary requirements. • Given a probability distribution, compute and use the mean and standard deviation. • Use probabilities to determine whether or not an event is significant.
5.2	• Understand the concepts related to binominal probability distributions. • Determine whether given distributions are binomial distributions. • Find probability values for a binomial distribution. • Compute and use the mean and standard deviation for a binomial distribution.

6.1	• Describe the characteristics of a standard normal distribution. • Find z scores corresponding to regions under the curve representing a standard normal distribution • Find the probability of some range of values in a continuous uniform distribution.
6.2	• Describe a normal distribution (not necessarily a standard normal distribution). • Find the probability of some range of values in a normal distribution. • Find x scores corresponding to regions under the curve representing a normal distribution. • Solve applications involving probabilities and corresponding x scores for a normal distribution.
6.3	• Develop the ability to describe a sampling distribution of a statistic. • Determine whether a statistic serves as a good estimator of the corresponding population parameter
6.4	• Check conditions for using a normal distribution for the distribution of sample means. • Describe what the central limit theorem states. • Use the central limit theorem to find the probability that a sample mean falls in a range of value • Identify the stated rule.
7.1	• Construct and interpret a confidence interval estimate of a population proportion. • Identify values and verify the requirements necessary to estimate a population proportion. • Determine the sample size necessary to estimate a population proportion.
7.2	• Identify values and verify the requirements necessary to estimate a population mean. • Construct and interpret a confidence interval estimate of a population mean. • Determine the sample size necessary to estimate a population mean.
7.3	• Identify values and verify the requirements necessary to estimate a population standard deviation • Construct and interpret a confidence interval estimate of a population standard deviation/variance • Determine the sample size necessary to estimate a population standard deviation or variance.
8.1	• Understand the concepts related to hypothesis testing. • Identify the null and alternative hypotheses when given some claim about a population parameter. • Calculate a test statistic. • Use P-values to state a final conclusion that addresses an original claim. • Use critical values to state a final conclusion that addresses an original claim. • Identify type I and type II errors that correspond to a given claim.
8.2	• Use sample data to conduct a formal hypothesis test of a claim about a population proportion.
8.3	• Use sample data to conduct a formal hypothesis test of a claim made about a population mean.
8.4	• Conduct a hypothesis test of a claim made about a population standard deviation or variance.
9.1	• Test claims and construct confidence interval estimates for the difference between two proportions

9.2	• Distinguish between two independent samples and two samples that are not independent. • Understand the concepts related to testing two means from independent samples. • Test a claim made about two means from independent populations.
9.3	• Conduct a test of a claim made about the mean of the differences between matched pairs.
9.4	• Conduct a hypothesis test of a claim made about two population standard deviations or variances.
10.1	• Understand the concepts related to correlation. • Understand the relationship between a scatter plot of paired data and the correlation coefficient. • Determine whether there is sufficient evidence to conclude that there is a linear correlation.
10.2	• Use paired sample data to find the equation of the regression line. • Find the best predicted value of a variable given the regression equation and a value. • Find the equation of a regression line and the best predicted value of a variable given a value. • Understand the concepts related to regression.
10.3	• Understand the concepts related to prediction intervals and variation. • Given a correlation coefficient, determine and interpret the coefficient of determination. • Use a given value of one variable to find a prediction interval for the other variable.
10.4	• Understand the concepts related to multiple regression. • Identify the predictor variables that result in the best multiple regression equation. • Determine whether a given multiple regression equation is suitable for making predictions.
10.5	• Understand the concepts related to nonlinear regression. • Compare the linear, quadratic, logarithmic, exponential, and power models for paired data.
12.1	• Conduct a hypothesis test of a claim made about two population standard deviations or variances. • Use one-way ANOVA to conduct a hypothesis test of equality of three or more population means. • Understand the concepts related to one-way ANOVA.
12.2	• Understand the concepts related to two-way ANOVA. • Interpret two-way ANOVA output to draw a conclusion for a hypothesis test for interaction/effects. • Use two-way ANOVA to conduct a hypothesis test for interaction and effects.

Unit Tests

1.4

4) The following data set represents Heather's average monthly charges (in $) for cable TV for the past 24 months.

105	125	110	98	102	115	110	123	118	101	95	128
110	105	122	107	118	107	117	125	116	110	101	107

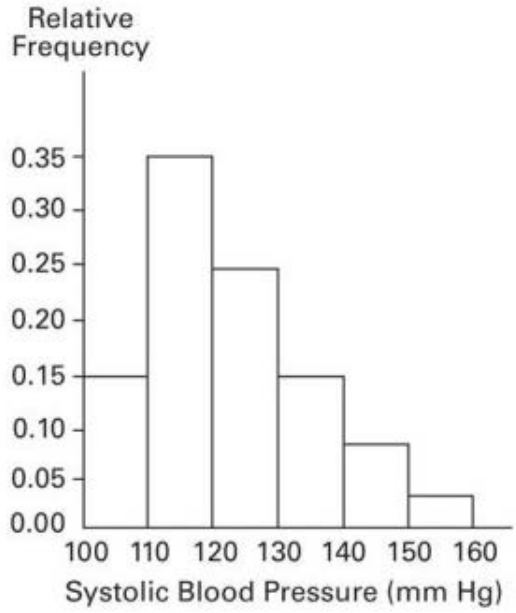
Construct a frequency distribution with 4 classes, a class width of $10, and a lower limit of 95 for class 1.

Charges $	Frequency

1.5

5) A nurse measured the blood pressure of each person who visited her clinic. Following is a relative-frequency histogram for the systolic blood pressure readings for those people aged between 25 and 40 years. The blood pressure readings were given to the nearest whole number. Approximately what percentage of the people aged 25–40 had a systolic blood pressure reading between 110 and 139 mm Hg inclusive?

Systolic Blood Pressure for People Aged 25–40 Years



1.6

6) The following data show the number of laps run by each participant in a marathon.

46 65 55 43 51 48 57 30 43 49 32 56

Which of these choices display the correct stemplot?

A)

3	0 2
4	3 6 8 9
4	1 3 5 6 7
6	5

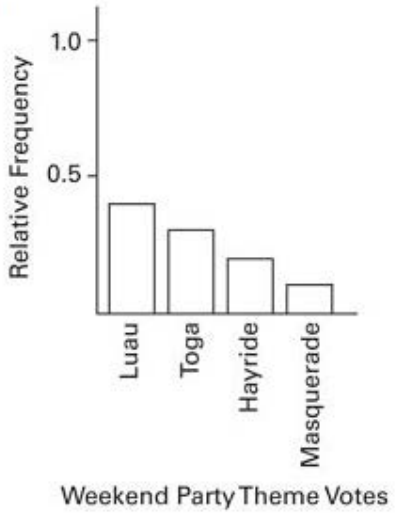
B)

3	0 2
4	3 3 6 8 9
5	1 5 6 7
6	5

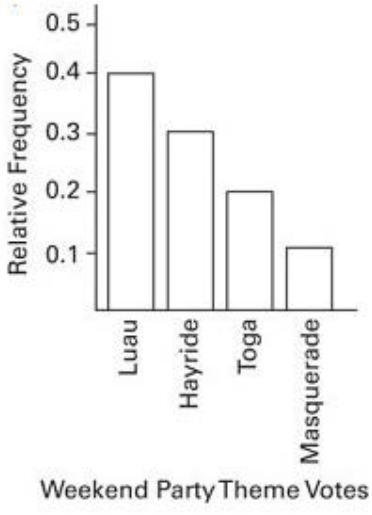
1.7

7) The Kappa Iota Sigma Fraternity polled its members on the weekend party theme. The vote was as follows: six for toga, four for hayride, eight for luau, and two for masquerade. Which of these choices display the correct Pareto chart?

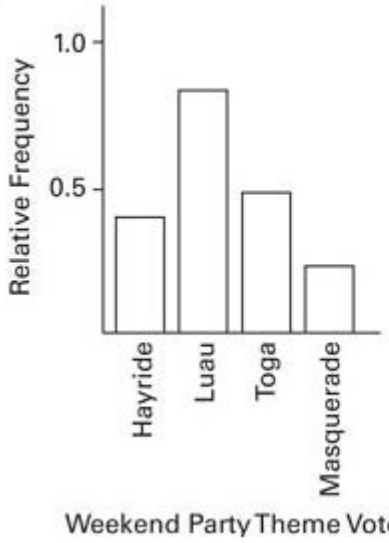
A)



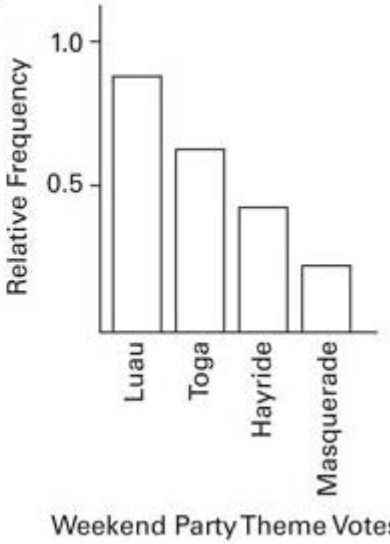
B)



C)



D)



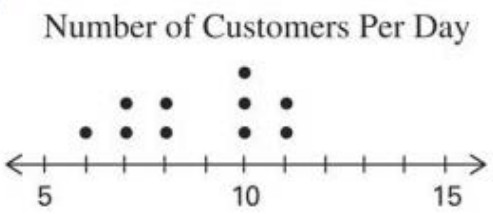
1.8

8) A store manager counts the number of customers who make a purchase in his store each day. The data are as follows.

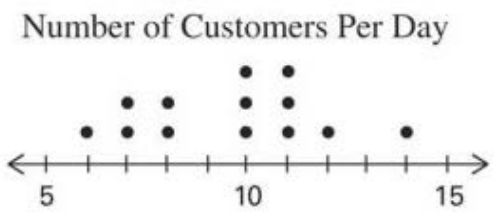
10 11 8 14 7 10 10 11 8 7

Which of these choices display the correct dotplot?

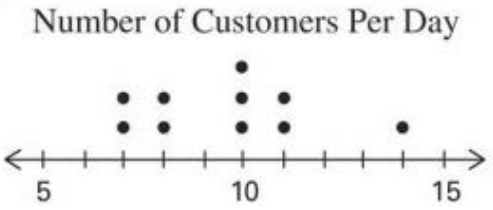
A)



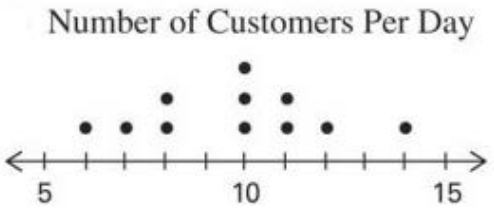
B)



C)

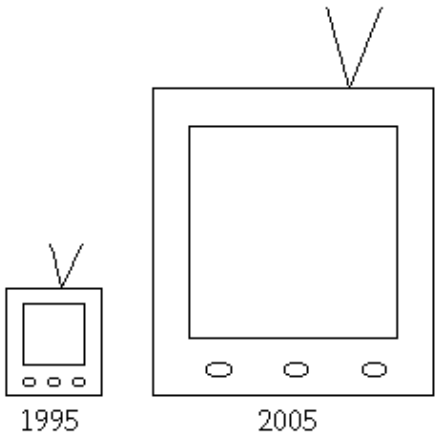


D)



1.9

9) A television manufacturer sold three times as many televisions in 2005 as it did in 1995. To illustrate this fact, the manufacturer draws a graph as shown below. The television on the right is three times as tall and three times as wide as the television on the left. Why is this graph misleading? What visual impression is created by the graph?



1.10

10) Listed below are the amounts of time (in months) that the employees of a restaurant have been working at the restaurant. Find the mean for the given sample data. Round your answer to one more decimal place than the original data values.

1 5 6 8 11 14 17 46 61 90 99 126 143 167

1.11

11) The number of vehicles passing through a bank drive –up line during each 15–minute period was recorded. The results are shown below. Find the median number of vehicles going through the line in a 15–minute period.

25	27	25	28
28	25	30	27
35	31	31	29
24	31	25	20
15	27	27	27

1.12

12) Last year, nine employees of an electronics company retired. Their ages (in years) at retirement are listed below. Find the mode(s) for the given sample data.

51 61 62 57 50 67 68 58 53

1.13

13) The speeds (in mph) of the cars passing a certain checkpoint are measured by radar. The results are shown below. Find the midrange.

44.3	41.4	42.7	40.6	43.1
40.5	44.8	42.0	44.3	42.1
43.4	42.0	40.6	43.4	41.4

1.14

14) Dave is a college student contemplating a possible career option. One factor that will influence his decision is the amount of money he is likely to make. He decides to look up the average salary of graduates in that profession. Which information would be more useful to him, the mean salary or the median salary. Why?

1.15

15) The prices (in dollars) of 12 electric smooth top ranges are listed below. Find the range for the given sample data.

865	1010	655	565	1465	1110
710	765	820	1310	555	1065

1.16

16) Find the standard deviation for the given sample data. Round your answer to one more decimal place than the original data.

18 18 18 9 15 5 14 5 15

1.17

17) The data set below consists of the scores of 15 students on a quiz. For this data set, which measure of variation do you think is more appropriate, the range or the standard deviation? Explain your thinking.

90	90	91	91	89
90	89	91	91	90
60	90	89	90	91

1.18

18) The race speeds (in minutes) for the top eight cars in a 200–mile race are listed below. Use the range rule of thumb to estimate the standard deviation. Round results to the nearest tenth.

181.0 180.6 189.2 182.2 175.6 180.0 177.9 181.8

1.19

Use the empirical rule to solve the problem.

19) The amount of Jen's monthly phone bill is normally distributed with a mean of $51 and a standard deviation of $10. What percentage of her phone bills are between $21 and $81?

1.20

20) The heights of the adults in one town have a mean of 67.1 inches and a standard deviation of 3.5 inches. What can you conclude from Chebyshev's theorem about the percentage of adults in the town whose heights are between 60.1 and 74.1 inches?

1.21

21) For a test score of 48.4 on a test having a mean of 66 and a standard deviation of 11, find the z score corresponding to the given value and use the z score to determine whether the value is significant. Consider a score to be significant if its z score is less than −2.00 or greater than 2.00. Round the z score to two decimal places if necessary.

1.22

22) Scores on the SAT test have a mean of 1518 and a standard deviation of 325. Scores on the ACT test have a mean of 21.1 and a standard deviation of 4.8. Which of the following choices is not true?

A) The ACT score of 17.0 is relatively better than the SAT score of 1490.
B) An SAT score of 1490 has a z score of −0.09.
C) The SAT score of 1490 is relatively better than the ACT score of 17.0.
D) An ACT score of 17.0 has a z score of −0.85.

1.23

23) In a data set with a range of 55.4 to 105.4 and 400 observations, there are 276 observations with values less than 86. Find the percentile for 86.

1.24

24) The test scores of 40 students are listed below. Find P_{56}.

30 35 43 44 47 48 54 55 56 57
59 62 63 65 66 68 69 69 71 72
72 73 74 76 77 77 78 79 80 81
81 82 83 85 89 92 93 94 97 98

1.25

25) The following are interval times (minutes) between eruptions of the Old Faithful geyser in Yellowstone National Park (based on U.S. National Park Service data).

81 81 86 87 89 92 93 94 95 96 97 98 98 101 101 106

Identify the five number summary that would be used to construct a boxplot of these data.

2.8

Determine whether the following is a probability distribution. If not, identify at least one requirement that is not satisfied.

8)

x	P(x)
1	0.037
2	0.200
3	0.444
4	0.296

2.9

9) Find the mean and the standard deviation for the given probability distribution. Round each answer to the nearest tenth.

x	P(x)
0	0.16
1	0.25
2	0.25
3	0.12
4	0.22

2.10

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

x(girls)	P(x)	x(girls)	P(x)	x(girls)	P(x)
0	0.000	5	0.122	10	0.061
1	0.001	6	0.183	11	0.022
2	0.006	7	0.209	12	0.006
3	0.022	8	0.183	13	0.001
4	0.061	9	0.122	14	0.000

10) Find the probability of selecting 10 or more girls. If 14 newborns are randomly selected, is 10 a significantly high number of girls?

2.11

11) In a survey of 300 college graduates, 47% 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? Round to the nearest thousandth.

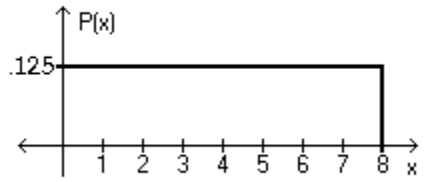
2.12

Determine if the outcome is significant. Consider as significant any result that differs from the mean by more than 2 standard deviations. That is, significant values are either less than $\mu - 2\sigma$ or greater than $\mu + 2\sigma$.

12) 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 611 consumers be a significantly high number of consumers who recognize the Dull Computer Company name? Justify your answer.

2.13

Using the following uniform density curve, answer the question.



13) What is the probability that the random variable has a value greater than 5?

2.14

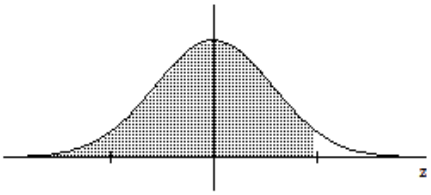
If z is a standard normal variable, find the probability.

14) The probability that z lies between −1.10 and −0.36. Round to four decimal places.

2.15

Find the indicated z score rounded to two decimal places. The graph depicts the standard normal distribution with mean 0 and standard deviation 1.

15) Shaded area is 0.9599.



2.16

Find the indicated value rounded to two decimal places.

16) $z_{0.01}$

2.17

Assume that X has a normal distribution, and find the indicated probability rounded to four decimal places.

17) The mean is $\mu = 15.2$ and the standard deviation is $\sigma = 0.9$.
Find the probability that X is greater than 16.1.

2.18

18) In one region, the September energy consumption levels for single −family homes are found to be normally distributed with a mean of 1050 kWh and a standard deviation of 218 kWh. For a randomly selected home, find the probability that the September energy consumption level is between 1100 kWh and 1225 kWh.

2.19

19) A bank's loan officer rates applicants for credit. The ratings are normally distributed with a mean of 200 and a standard deviation of 50. Find P_{60}, the score which separates the lower 60% from the top 40%. Round to the nearest whole.

2.20

20) The annual precipitation amounts in a certain mountain range are normally distributed with a mean of 109 inches, and a standard deviation of 10 inches. What is the probability that the mean annual precipitation during 25 randomly picked years will be less than 111.8 inches? Round to four decimal places.

3.1

1) Find the critical value $z_{\alpha/2}$ that corresponds to a 91% confidence level. Round your answer to two decimal places.

3.2

2) Assume that a sample is used to estimate a population proportion p. Find the margin of error E that corresponds to the given statistics and confidence level. Round your answer to four decimal places.

95% confidence; $n = 2388$, $x = 1672$

3.3

3) Of 380 randomly selected medical students, 21 said that they planned to work in a rural community. Find a 95% confidence interval for the true proportion of all medical students who plan to work in a rural community. Round endpoints to four decimal places. Then write a sentence explaining what this confidence interval means.

3.4

4) Use the given data to find the minimum sample size required to estimate the population proportion.

Margin of error: 0.028; confidence level: 99%; $\hat{p}$ unknown

3.5

5) Which critical value is appropriate for a 99% confidence level where $n = 17$; σ is unknown and the population appears to be normally distributed?

A) $t_{\alpha/2} = 2.898$ B) $z_{\alpha/2} = 2.583$ C) $z_{\alpha/2} = 2.567$ D) $t_{\alpha/2} = 2.921$

3.6

6) 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.

A laboratory tested twelve chicken eggs and found that the mean amount of cholesterol was 185 milligrams with $s = 17.6$ milligrams. Construct a 95% confidence interval for the true mean cholesterol content of all such eggs. Round endpoints to one decimal place. Then write a sentence explaining what this confidence interval means.

3.7

Use the given information to find the minimum sample size required to estimate an unknown population mean μ.

7) How many commuters must be randomly selected to estimate the mean driving time of Chicago commuters? We want 90% confidence that the sample mean is within 4 minutes of the population mean, and the population standard deviation is known to be 12 minutes.

3.8

8) Find the critical value χ^2_R corresponding to a sample size of 19 and a confidence level of 99 percent. Round to three decimal places.

3.9

9) To find the standard deviation of the diameter of wooden dowels, the manufacturer measures 19 randomly selected dowels and finds the standard deviation of the sample to be $s = 0.16$ cm. Find the 95% confidence interval for the population standard deviation σ. Round endpoints to two decimal places. Then write a sentence explaining what this confidence interval means.

3.10

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.

10) A cereal company claims that the mean weight of the cereal in its packets is at least 14 oz.

3.11

Assume that a hypothesis test of the given claim will be conducted. Identify the type I error for the test.

11) A medical researcher claims that 17% 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 equal to 17% when that percentage is actually 17%.

B) Fail to reject the claim that the percentage of children who suffer from the disorder is equal to 17% when that percentage is actually different from 17%.

C) Fail to reject the claim that the percentage of children who suffer from the disorder is equal to 17% when that percentage is actually 17%.

D) Reject the claim that the percentage of children who suffer from the disorder is different from 17% when that percentage really is different from 17%.

3.12

Identify which distribution is used for the given claim. The find the value of the test statistic using the appropriate formula. Round to two decimal places.

12) A claim is made that the proportion of children who play sports is less than 0.5, and the sample statistics include n = 1906 subjects with 30% saying that they play a sport.

3.13

13) Use the given information to find the P-value rounded to four decimal places. Also, use a 0.05 significance level and state the conclusion about the null hypothesis (reject the null hypothesis or fail to reject the null hypothesis). With $H_1 : p \neq 0.377$ the test statistic is $z = 3.06$.

3.14

Identify the null hypothesis, alternative hypothesis, test statistic (rounded to two decimal places), P–value (rounded to four decimal places), conclusion about the null hypothesis, and final conclusion that addresses the original claim.

14) A poll of 1068 adult Americans reveals that 48% of the voters surveyed prefer the Democratic candidate for the presidency. At the 0.05 level of significance, test the claim that at least half of all voters prefer the Democrat.

3.15

Identify the null hypothesis, alternative hypothesis, test statistic (rounded to three decimal places), P–value (rounded to four decimal places), conclusion about the null hypothesis, and final conclusion that addresses the original claim.

15) A public bus company official claims that the mean waiting time for bus number 14 during peak hours is less than 10 minutes. Karen took bus number 14 during peak hours on 18 different occasions. Her mean waiting time was 7.7 minutes with a standard deviation of 2 minutes. At the 0.01 significance level, test the claim that the mean waiting time is less than 10 minutes.

3.16

Identify the null hypothesis, alternative hypothesis, test statistic (rounded to three decimal places), P–value (rounded to four decimal places), conclusion about the null hypothesis, and final conclusion that addresses the original claim.

16) In one town, monthly incomes for men with college degrees are found to have a standard deviation of $650. Use a 0.01 significance level to test the claim that for men without college degrees in that town, incomes have a higher standard deviation. A random sample of 22 men without college degrees resulted in incomes with a standard deviation of $903.

3.17

Construct and interpret 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. Round the endpoints of the interval to three decimal places.

17) 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. Construct a 95% confidence interval for the difference between the population proportions $p_1 - p_2$. Then write a sentence interpreting the confidence interval.

3.18

Test the indicated claim about the means of two populations. Assume that the two samples are independent simple random samples selected from normally distributed populations. Use the P–value method. Identify the null hypothesis, alternative hypothesis, test statistic (rounded to three decimal places), P–value (rounded to four decimal places), conclusion about the null hypothesis, and final conclusion that addresses the original claim.

18) A researcher was interested in comparing the amount of time (in hours) spent watching television by women and by men. Independent simple random samples of 14 women and 17 men were selected, and each person was asked how many hours he or she had watched television during the previous week. The summary statistics are as follows.

Women	Men
$\bar{x}_1 = 12.5$ hr	$\bar{x}_2 = 13.8$ hr
$s_1 = 3.9$ hr	$s_2 = 5.2$ hr
$n_1 = 14$	$n_2 = 17$

Use a 0.05 significance level to test the claim that the mean amount of time spent watching television by women is smaller than the mean amount of time spent watching television by men.

3.19

Construct and interpret 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.

19) The table below shows the weights of 9 subjects before and after following a particular diet for two months.

Subject	A	B	C	D	E	F	G	H	I
Before	168	180	157	132	202	124	190	210	171
After	162	178	145	125	171	126	180	195	163

Construct a 99% confidence interval for the mean difference of the "before" minus "after" weights. Round endpoints to one decimal place. Then write a sentence interpreting the confidence interval.

3.20

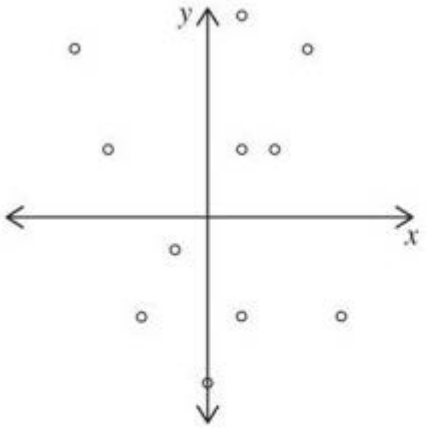
Test the indicated claim about the variances or standard deviations of two populations. Assume that both samples are independent simple random samples from populations having normal distributions. Identify the null hypothesis, alternative hypothesis, test statistic (rounded to three decimal places), P–value (rounded to four decimal places), conclusion about the null hypothesis, and final conclusion that addresses the original claim.

20) A random sample of 16 women resulted in blood pressure levels with a standard deviation of 23 mm Hg. A random sample of 17 men resulted in blood pressure levels with a standard deviation of 19.2 mm Hg. Use a 0.05 significance level to test the claim that blood pressure levels for women vary more than blood pressure levels for men.

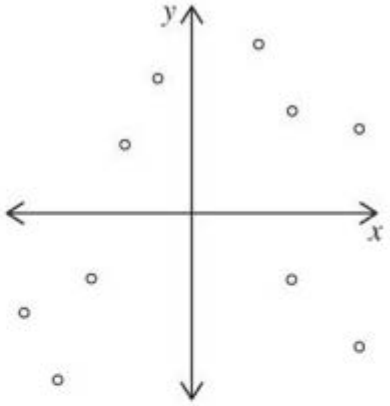
4.1

1) Determine which plot shows the strongest linear correlation.

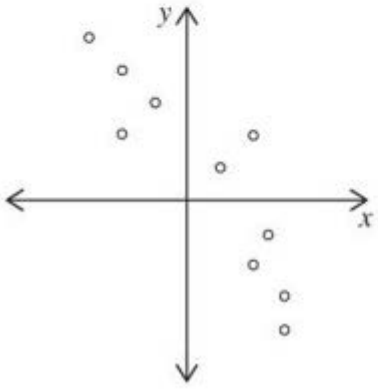
A)



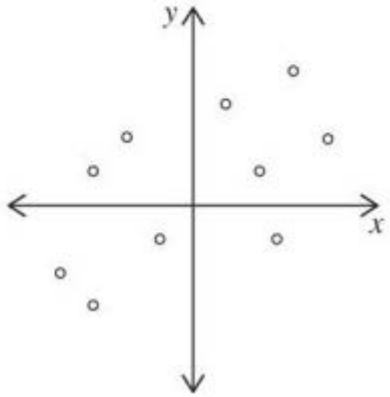
B)



C)



D)



4.2

2) The paired data below consist of the temperatures on randomly chosen days and the amount a certain kind of plant grew (in millimeters). Find the value of the linear correlation coefficient r (rounded to three decimal places). Then find the P-value (rounded to three decimal places), and, using a 0.01 significance level, determine whether there is sufficient evidence to support a claim of a linear correlation between the two variables.

Temperature	62	76	50	51	71	46	51	44	79
Growth	36	39	50	13	33	33	17	6	16

4.3

3) Two different tests are designed to measure employee productivity and dexterity. Several employees are randomly selected and tested with these results. Use the given data to find the equation of the regression line. Round the final values to three significant digits, if necessary. Then use that equation to predict the dexterity of an employee with a productivity level of 32. Is that predicted value the best predicted value? Why or why not?

Productivity (x)	23	25	28	21	21	25	26	30	34	36
Dexterity (y)	49	53	59	42	47	53	55	63	67	75

4.4

4) The paired data below consists of test scores and hours of preparation for 5 randomly selected students. The equation of the regression line is $\hat{y} = 44.845 + 3.524x$ and the margin of error for a 99% confidence level is approximately E = 35. Find the 99% prediction interval for the test score of a person who spent 7 hours preparing for the test. Round the endpoints of the interval to the nearest whole number. Then write a sentence explaining what that prediction interval means.

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

4.5

5) Find the coefficient of determination, given that the value of the linear correlation coefficient, r, for a collection of paired data is 0.738. Then find the percentage of the total variation that can be explained by the linear relationship between the two variables.

4.6

6) Below are performance, attitude, and adaptablity ratings of employees as given by their managers.

Performance:	59	63	65	69	58	77	76	69	70	64
Attitude:	72	67	78	82	75	87	92	83	87	78
Adaptability:	50	52	54	60	46	67	66	59	62	55

Use the table below to determine which regression equation is best for predicting performance level. Justify your answer.

Predictor Variables	P-value	R^2	Adj R^2	Regression Equation
ATT/ADAPT	0.000	0.973	0.965	PERF = 14.1 + 0.0138ATT + 0.907ADAPT
ATT	0.001	0.745	0.499	PERF = 8.57 + 0.730ATT
ADAPT	0.000	0.973	0.940	PERF = 14.4 + 0.921ADAPT

4.7

7) The table below shows the population of a city (in millions) in each year during the period 2010–2015. Different non-linear regression models and their corresponding coefficients of determination are given below. (Years since 2009 was used as the independent variable.) Identify the model that best fits the data.

Year	2010	2011	2012	2013	2014	2015
Population(millions)	1.08	1.37	1.68	2.19	2.73	3.34

Linear: $y = 0.476 + 0.454x$; $r^2 = 0.9789$
Quadratic: $y = 0.045x^2 + 0.139x + 0.896$; $R^2 = 0.9994$
Logarithmic: $y = 0.755 + 1.19\ln x$; $r^2 = 0.8507$
Exponential: $y = 0.863(1.26)^x$; $r^2 = 0.9988$
Power: $y = 0.967x^{0.624}$; $r^2 = 0.9394$

4.8

8) The data below represent the weight losses for people on three different exercise programs. Test the claim that the samples come from populations with the same mean. Assume that the populations are normally distributed with the same variance.

Exercise A	Exercise B	Exercise C
2.5	5.8	4.3
8.8	4.9	6.2
7.3	1.1	5.8
9.8	7.8	8.1
5.1	1.2	7.9

At the 1% significance level, does it appear that a difference exists in the true mean weight loss produced by the three exercise programs? Give both the test statistic and P-value to four decimal places.

4.9

9) 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 test statistics and P-values to four decimal places.

	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

Attachment 2 – Results Separated by Performance Indicator