John Climent
Office Location: AS 316
Office Phone: 410-287-6060 (Ext. 386)
Home Phone: 410-398-3733 (Between 9:00 A.M. & 10:00 P.M.)
E-mail Address: jcliment@cecilcc.edu

TIME REQUIRED TO PASS THIS COURSE:

Typically students have to spend 15 to 20 hours per week to be successful in this course. Don't take this statement lightly. It comes from students who have taken this course. Ask fellow students who have taken this course and they will confirm this statement. Here is a typical 16 to 20 hour week: 4 hours for class, 3 to 4 hours for ALS computer certificates, 3 hours for homework, 4 to 5 hours for take-home test, 2 to 4 hours for other activities (i.e., project, experiments, Minitab, in-lab tests, etc.).

PREREQUISITES:

The prerequisite for this course is a demonstration of a satisfactory level of Intermediate Algebra as shown by a grade of C or better in MAT 093 (Intermediate Algebra) or appropriate placement on the Mathematics Skills Assessment, as determined by the Math Department.

TEXT:

Required: Discovering Statistics, An Adventure in Problem Solving by James S. Hawkes.
Required: Adventure Learning Systems: Statistics - software to accompany text. (Note: software cannot be purchased used.)

COURSE OUTLINE:

Testing Unit	Chapters Covered	Evaluation
1. Descriptive Statistics	1. Statistics and Problem Solving 2. Data and Problem Solving 3. Displaying Data 4. Describing Data from One Variable 5. Discovering Relationships	Homework Regression Project Test 1 – Take-Home Portion Test 1 – Math Lab Portion ALS Certificates
2. Probability Distributions	6. Probability, Randomness, and Uncertainty 7. Probability Distributions, Information about the Future 8. Continuous Random Variables 9. Samples and Sampling Distributions	Homework Test 2– Take-Home Portion Test 2– Math Lab Portion ALS Certificates
3. Inferential Statistics	10. Estimating Means and Proportions 11. Hypothesis Testing (Note: selected sections of chapters 12-15 may be covered if time permits.)	Homework Inferential Statistics Project Test 3– Take-Home Portion Test 3– Math Lab Portion ALS Certificates

GADING:

The course is divided up into three units (Descriptive Statistics, Probability Distributions, and Inferential Statistics). Your grade in the course will be computed as the weighted average of your grade on these three units and final exam is as follows:

Course Grade = 0.30 · Unit 1 Grade + 0.30 · Unit 2 Grade + 0.25 · Unit 3 Grade+ 0.15 · Final + Attendance Bonus

Note: although the final exam is cumulative, its emphasis will be on Unit 3.
See attendance policy for the attendance bonus details.

The percentage weightings within each unit are shown in the table below.			
Unit	**1. Descriptive Statistics**	**2. Probability Distributions**	**3. Inferential Statistics**
Homework	10%	10%	10%
Project	15%	NA	15%
Take-Home Test	30%	30%	45%
Math Lab Test	25%	30&	NA
ALS Certificates	20%	30%	30%

Letter Grade Computation: A = 90%-100%; B = 80%-89%; C = 70-79%; D = 60-69%; F = 0-59%

The administrative withdrawal grade, M, may be assigned in accordance with the college policy and procedure. Since the "M" grade can have the same effect on a student as the "W" grade, any student who earns an "M" grade, but does not want one, must notify the instructor in writing. Such students will be given F's instead of M's.

PROJECTS:

There are two projects in this course. One on regression analysis, which will be due after chapter 5, and the other is on inferential statistics, which will be due at the end of the semester. Detailed written instructions for each project will be given later. Projects may be presented both orally and in writing. Students may work in small groups (up to 3 students) on a project. For most students the write up of the project only takes a few hours. Generally the hardest part is figuring out what data to gather and then gathering it. The gathering of data can be done throughout the semester, but you will need help from your instructor in figuring out what data you need to gather. This is especially true for the project due at the end of the semester. Do not begin gathering the data for your project before discussing it with your instructor. You may make an appointment with your instructor at any time to discuss the project and the data that you will need to gather. **Do not begin your project or collect data for it without first getting the approval of your instructor. Some students have done this in the past only to find that the data they have gathered is useless.**

HOMEWORK:

Description: You will be given an assignment sheet for the semester. Generally, you will be expected to do the assignments for the work covered in each class session prior to the next class session. The beginning of each class session will be set aside to answer questions and go over homework from the assignment given during the previous class. You may be called upon at any time to put homework problems on the blackboard.

Grading the Homework: Homework from the previous week will be due at the very beginning of first class each week. Your homework will examine for its completeness, correctness and effort. **Just copying answers out of the back of the text is insufficient** (except in a few specific instances). Homework will be assessed not graded and a scale of 0 – 2 shall be used.

Importance of Homework: Almost every individual needs a lot of practice to learn mathematics. It is not one of those subjects that one learns just by reading or listening. You must practice it to learn it well. Your instructor's lectures and the explanations in your text are only intended to help make your learning and understanding of the subject easier. You must do your part. Unless you practice math, you will most likely never learn it. Homework is where you get the most practice and learn the most. *Some individuals may need to do more than the assigned problems to learn the subject well.* Since homework plays such an important part in learning mathematics, a large portion of your grade is determined by it. It is essentially worth a full letter grade in determining your course grade.

UNIT TESTS – MATH LAB PORTION (Test 1, Test 2 & the Final):

The Math Lab Portion of Tests 1 & 2 will be given in the Math Lab. Test 1 will be given after chapter 5,and Test 2 will be given after chapter 9. The Math Lab Portion of the Final will either be in the Lab or on the day during the last week of the semester set aside for the final exam. No make up finals will be given unless prior arrangements are made with your instructor.

UNIT TESTS – TAKE HOME PORTION (Test 1, Test 2 & the Final):

You will be given a take-home test on each unit. They will be given in segments usually corresponding to chapters in the book. Each segment will have its own due date. The grade for the take home tests will be the average of the accumulated points for all segments.

ADVENTURE LEARNING SYSTEMS: Statistics:

The **Adventure Learning Systems: Statistics** software that was created to accompany the text is mandatory. Your assignment sheets will specify when you should attempt each of the **ALS** lessons. You need to certify that you have successfully completed each lesson. Your **ALS** grade for each unit will simply be the average of your scores for each **ALS** lesson assigned for that unit.

ATTENDANCE:

You are expected to attend all classes. Perfect attendance = 5% bonus in final grade, exactly one absence = 3% bonus in final grade, exactly two absences = 1% bonus in final grade, and 3 or more absences = 0% bonus in final grade. Lateness for class will generally be considered half an absence, unless you show up at the end of class. There will be no excused absences. If you are a student who must depend on the attendance bonus, then you should arrange to attend all classes on time.

CALCULATORS:

Calculators are required for all assignments, tests and exams. It is recommended that you purchase one that has statistical functions and that you learn how to use it. See your instructor for help in choosing and learning how to use your calculator. Many students

have used graphing calculators in previous math courses. The Texas Instruments **TI-83** graphing calculator **is the best calculator** for this course. It has a statistical capability that can do nearly all of the computations required in this course effortlessly. **If you are buying a graphical calculator for use in this course, don't let some sales person talk you into a different brand or model.** The older version of this calculator, the TI-82, also can do many of the computations required in this course. The statistics functions on these two calculators are fairly easy to learn how to use. Your instructor can help you with this. It is not recommended that you purchase such an expensive calculator just for this course, but since many students already have them, it is highly recommended that you learn how to fully use their statistical capabilities. Use of one of the calculators may be more convenient for those students who cannot come to the college to do the work on their quizzes and test on the computers available there. It is possible to rent one from the College for $25 per semester. See Patricia Voigt in the Math Lab if you wish to rent one.

PERSONAL COMPUTERS:

Minitab: A statistical software package called Minitab will be used extensively in this course. It is available on the computers in the classroom, the Math Lab and other locations throughput the campus. Your instructor will give you help in learning how to use Minitab. It is a fairly easy piece of software to use. Minitab is copyright protected, so you may not take it home with you, nor may you copy it. It is possible to rent Minitab for about $26 per semester (5 months). The rental details can be found on the following web site: http://www.minitab.com/products/semester_rental/INDEX.HTM. You can also download a one-month trial version of Minitab from that web site. Newer versions of spread sheeting packages, such as Microsoft Excel, have increased their ability to do statistics. Students who are familiar with these packages may use them for some and possible all of the computer assignment, but there are two downsides or risks to using them. Sometimes they give you incorrect answers, because they are not doing what you think they are doing. In addition, there is no support from the personnel in the Math Lab for these programs. ***Note: not everyone in the Math Lab can give you help with Minitab, so try to schedule your sessions there when help is available.***

EXEMPTIONS FROM CLASS REQUIREMENTS:

If there is any reason why you cannot meet any of the requirements of this course, notify your instructor immediately. Exemptions or alternative requirements are always possible. They will be given at the discretion of your instructor. Requests should be submitted in writing.

TAKE-HOME TEST PROCEDURE AND REQUIREMENTS:

There are three rules you are asked to follow on the take-home exams. You may not copy your work or answers from someone else, you may not have another person do the problems for you and you may not compare answers with those of another student. This does not mean that you cannot seek help, or work with other students. Just the opposite is true. If you find you do not know how to do a certain problem, you are encourage to go to the Math Lab and ask them for help with a problem that is similar to the one you are working on. You may not ask them to do the actual problems for you, that is considered cheating, but you may ask them to show you how to do similar problems. You may work with other students doing problems similar to the exam problems, but when you do the actual exam problems, you must do them by yourself. You may also use your text or any other book for the same purpose. Your instructor has chosen to give you take-home exams because they relieve a lot of the anxiety concerning math testing and because they are a very good learning tool when taken properly. However, if you wait until the last minute to do the test and rush through it, you are losing its value as a learning tool. In general, the take-home exams may be a little more challenging and a little longer than exams normally given in class.

IN-LAB TEST PROCEDURE AND REQUIREMENTS:

All In-Lab tests are to be taken in the area designated for testing. Unless specified otherwise, you are not allowed to have any materials in the test area other than the test itself, scrap paper, writing materials, and calculators. You are not allowed to talk with other students and you must follow any instructions given to you by the Math Lab personnel. There is no limit on how much time you may spend on a test taken in the Math Lab; however, the Math Lab will not stay open past its posted hours for you to finish a test. It is your responsibility to make sure that you begin your test in plenty of time to finish it. If you are not sure you have enough time to complete a test, you may take it one page at a time. When you leave, you must hand in your test, all scrap paper, and all written materials used in taking the test. Failure to follow the procedures outlined above may result in your receiving a failing grade in the course. If for some reason, you find you are having trouble taking a test in the Math Lab, notify the Lab personnel immediately. They are authorized to accommodate reasonable requests. In the case of disputes or unforeseen circumstances, such as the Lab being closed during its posted hours, leave your instructor a note in his mailbox or on his voice mail. This will ensure that you are not held responsible for circumstances that are not caused be your actions.

CHEATING:

Copying another person's work or allowing your work to be copied by someone else is considered cheating. If your instructor has determined that you have cheated on a test, assignment, project, etc., you will receive a zero on that particular assignment. Any subsequent cheating will result in failure for the course. It is very tempting to cheat on take-home exams. It is sometimes very easy for your instructor to detect such cheating. Your instructor has an easy method for detecting when students have copied answers from his answers, so if the opportunity to copy answers from an answer key presents itself to you, be aware that you will most likely be caught and penalized. As mentioned earlier, you may work together on take-home exams by doing problems similar to the exam problems, but you may not work together on the exact exam problems, nor may you compare or copy answers to exam problems. If

your instructor suspects it is possible that you have cheated, you will be required meet with him at his convenience to explain how you got your answers. Note: exams with similar wrong answers are automatically suspected of cheating and in most such instances you will be asked to explain how you got your answers. Often there are good reasons why two papers get similar wrong answers but sometimes the reason is that the students cheated. If you are asked to explain why your paper is similar to someone else's paper you shouldn't have anything to worry about, unless you have actually cheated. Failure to meet with your instructor concerning your exam shall be considered an admission of guilt concerning cheating.

CONTACTING YOUR INSTRUCTOR:

At Cecil: You may call your instructor anytime at 410-287-6060 (Extension 386). If he is not there you may leave a message on the voice mail system. Messages are only checked two or three times a week.

At Home: You can also your instructor at home (410-398-3733). Do not call before 9 in the morning or after 10 at night. If you get an answering machine (6 or 7 rings) leave a message that includes the phone number(s) where you can be reached and how late he can return your call.

By E-mail: jcliment@cecilcc.edu

MAT 127 SPRING 2002 ASSIGNMENT SHEET – REVISED 11/12/2002 JOHN CLIMENT

Text: Discovering Statistics, An Adventure in Problem Solving by James S. Hawkes
Software: Adventure Learning Systems: Statistics

UNIT 1 ASSIGNMENT

Day	Date	Section	Work Assigned. It's due the next class and assessed the first class in the following week.
Thur.	8/29	1.1	**A:** 1,2
		1.2-1.5	**B:** 1,3,5
		Data	Gather Data for Pennies Data Set. (See Weight of Pennies Assignment.)
Thur.	9/5	2.1-2.4	**A:** 1,3,5,7
		2.5-2.8	**B:** 1,3,5,7,11
		ALS	Certify on Adventures in Learning Systems Lesson 1.1: Level of Measurement.
Mon.	9/9	3.1-3.6	**A:** 1,3,7,9
		Data	Draw a Histogram of your Pennies Data Set (See Pennies Assignment).
		Data	Gather Data for Peanut Data Set. (See Weight of Peanuts Assignment.)
Thur.	9/12	3.7-3.10	**B:** 1,3
		Data	Draw a Time-Series graph of the Frequency of Pennies for your data.
		ALS	Certify on Adventures in Learning Systems Lesson 2.1: Understanding and Interpreting Graphs.
Mon.	9/16	4.1-4.3	**A:** 1,7,9,11
		4.4	**B:** 1,7,9
		Data	Calculate Descriptive Statistics for the **Sample** of Pennies. (See Pennies Assignment).
		ALS	Certify on Adventures in Learning Systems Lesson 2.2: Descriptive Statistics.
Thur.	9/19	4.5	**C:** 5,7,9,10,11
		4.6-4.10	**D:** 1,3,5,6,7
		4.11	**E:** 1,3
		Data	Draw Box Plots of the "New" & "Old" Pennies. (See Pennies Assignment.)
		ALS	Certify on Adventures in Learning Systems Lesson 2.3: Constructing Samples.
		ALS	Certify on Adventures in Learning Systems Lesson 2.4: Using Descriptive Statistics.
Mon.	9/23	5.1-5.4	**A:** 1,3,5
		Data	Draw a Scatter Plot, etc. of the Pennies Data. (See Pennies Assignment).
		ALS	Certify on Adventures in Learning Systems Lesson 9.1: Fitting a Linear Model.
Thur.	9/26	5.5-5.6	**B:** 1,3,5,7,9
		Data	Gather Data for Project 1.
Mon.	9/30	5.7-5.13	**C:** 1,5,7
Thur.	10/3	5.14	**D:** 1,3,5

UNIT 2 ASSIGNMENT

Mon.	10/7	6.1-6.5	**A:** 1,3,5,7,9,11,13
Thur.	10/10	Projects	**Student Projects Due:** All Students Present their First Projects During this Class.
Mon.	10/14	6.6-6.10	**B:** 1,3,4,5,6,7,9,11,13,15
Thur.	10/17	6.11-6.12	**B:** 17, 18, 19, 20,21,22 **More Problems:** 1,2,3,4
Mon.	10/21	7.1-7.3	**A:** 1,3,7,9,11
		7.4-7.7	**B:** 1,3,5,7
		ALS	Certify on Adventures in Learning Systems Lesson 3.7: Discrete Random Variables.
Thur.	10/24	7.8	**C:** 1,3,5,7,9,11
		ALS	Certify on Adventures in Learning Systems Lesson 3.5: Binomial Word Problems.
Mon.	10/28	7.9	**D:** 1,3,5,7
		7.10	**E:** 1,3
		ALS	Certify on Adventures in Learning Systems Lesson 3.6: Poisson Word Problems.
		ALS	Certify on Adventures in Learning Systems Lesson 3.8: Hypergeometric Word Problems.
Thur.	10/31	8.1-8.2	**A:** 1,3
		8.3-8.4	**B:** 1,2,3,4,5,7,9,11,13,15,17,19
		8.5	**C:** 1,3,5,7,9
		ALS	Certify on Adventures in Learning Systems Lesson 4.1: The Standard Normal.
		ALS	Certify on Adventures in Learning Systems Lesson 4.2: Normal Distribution Word Problems.

Mon.	11/4	9.1-9.4	**A:** 1,3,5
		9.5-9.7	**B:** 1,3,4,5,6,7,9,11
		ALS	View on Adventures in Learning Systems Lesson 5.1: Sampling Distributions Simulation.
		ALS	Certify on Adventures in Learning Systems Lesson 5.2: Sampling Distributions (Means).
Thur.	11/7	9.8	**C:** 1,3,5,7,9,11
		9.9	**D:** 1,3,5
		ALS	Certify on Adventures in Learning Systems Lesson 5.3: Sampling Distributions (Proportions).
UNIT 3 ASSIGNMENT			
Mon.	11/11	10.1-10.4	**A:** 1,3,5,7,9,11
		10.5	**B:** 1,3,5,7,9
		10.6	**C:** 1,3
		ALS	Certify on Adventures in Learning Systems Lesson 4.3: Find the value of z.
		ALS	Certify on Adventures in Learning Systems Lesson 4.4: Find the value of t.
		ALS	Certify on Adventures in Learning Systems Lesson 6.1: Estimation (Means).
		ALS	Certify on Adventures in Learning Systems Lesson 6.2: Estimation (Means) Small Samples.
Thur.	11/14	**No Class – Attending National Conference**	
Mon.	11/18	10.8-10-8	**D:** 1,3,5
		10.9	**E:** 1,3
		ALS	Certify on Adventures in Learning Systems Lesson 6.3: Estimation (Proportions).
Thur.	11/21	11.1-11.3	**A:** 1
		11.4-11.6	**B Large Sampler:** 1,3,5,7,9,15
		ALS	Certify on Adventures in Learning Systems Lesson 7.2: Hypothesis Testing Means (z Value).
		11.4-11.6	**B Small Samples:** 1,3,5,7,9,13
		ALS	Certify on Adventures in Learning Systems Lesson 7.3: Hypothesis Testing Means (t Value).
		11.4-11.6	**B P-Values:** 1,3,5,7
		ALS	Certify on Adventures in Learning Systems Lesson 7.1: Hypothesis Testing Means (P Value).
Mon.	11/25	11.7	**C:** 1,3,5,9,11,13
		ALS	Certify on Adventures in Learning Systems Lesson 7.5: Hypothesis Testing Props. (z Value).
		ALS	Certify on Adventures in Learning Systems Lesson 7.4: Hypothesis Testing Props. (P Value).
		11.8*	**D:** 1,3,5
		ALS*	Certify on Adventures in Learning Systems Lesson 7.8: Hypothesis Testing Pop. Variance.
Thur.	11/28	**Thanksgiving – No Class.**	
Mon.	12/2	10.9	**E Large Sample:** 1,3,5,7
		ALS	Certify on Adventures in Learning Systems Lesson 8.1: Comparing Two Means (Large Sample).
		10.9	**E Small Sample:** 1,3,5,7
		ALS	Certify on Adventures in Learning Systems Lesson 8.2: Comparing Two Means (Small Sample).
		11.10	**F:** 1,3,5
		ALS	Certify on Adventures in Learning Systems Lesson 8.3: Comparing Two Means (Dependent).
		11.11	**G:** 1,3,5
		ALS	Certify on Adventures in Learning Systems Lesson 8.4: Comparing Two Proportions (Large).
Thur.	12/5	Review	**Review**
Mon.	12/9	Projects	**Student Projects Due:** All Students Present their Second Projects During this Class.
Items with a "*" may not be covered. Details will be given in class.			

Page 1 Test 1 – Lab MAT 127 Climent Name: ________________________________

Directions: 1. Show all work on these sheets. **Grade: ________**

2. **Calculators are permitted.**
3. **Computers are permitted and Minitab handouts may be used.**
4. **You may use hand written notes on a sheet of 8.5" by 11" sheet of paper.**
5. **You must hand in all written materials including your notes and scrap paper.**
6. **You must sign in and out for this test.**
7. **You may take this test one page at a time.**

1. (20 points) There are 42 students (8 men & 34 women) in the three sections of statistics this semester. The back-to-back stem plot on the right gives the points earned (maximum possible 70) on the third portion of the take-home test for the men and women enrolled. For all 42 students, find the quantities asked for and where indicated supply the appropriate statistical symbol for the quantity.

Leaves for Men	Stem	Leaves for Women
7	**3b**	
1	**4a**	4
7	**4b**	5 6 8 9
4	**5a**	1 2 2 3
5	**5b**	5 5 6 6 7 8 8 8 9
4 3	**6a**	0 1 1 2 3 3 3 3 3
8 5	**6b**	5 6 6 7 7 8

Round off to the nearest hundredth where necessary.

Symbol (1 point)	Value (2 points)	Quantity Rounded to Nearest **Hundredth** where necessary.
___________	___________	The mean grade.
___________	___________	The median grade.
Omit	___________	The 12% trimmed mean of the grades.
___________	___________	The variance of the grades.
___________	___________	The standard deviation of the grades.
Omit	___________	The mean absolute deviation of the grades.
Omit	___________	The mode of the grades.
Omit	___________	P_{85} for the grades.

2. (8 points) In the space below draw a box plot for the test grades from the previous problem. Use the number line provided, and label the value of the quartiles on the top of your box.

|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|
35 40 45 50 55 60 65
70

3. (6 points) The average daily temperature in Elkton last week was 67.0˚F. The average daily temperatures from Monday through Friday were 74.2˚F, 63.8˚F, 65.1˚F, 62.0˚F, & 67.5˚F.

________________ What is the level of this measurement?

________________ The average daily temperature for Saturday and Sunday were equal. What was their value?

4. (10 points) Based on the box plot shown find what is asked for. **Note: if it is not possible to answer one of these questions from the information given, simply write "More Info" in the blank for your answer.**

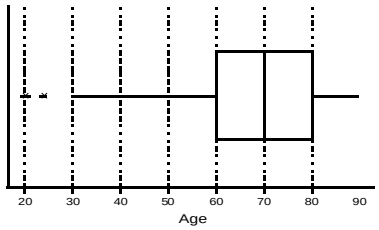
_____ What is the value of the mean?

_____ What is the value of the median?

_____ What is value of the interquartile range?

_____ How many outliers are there?

_____ What is the value of the range?



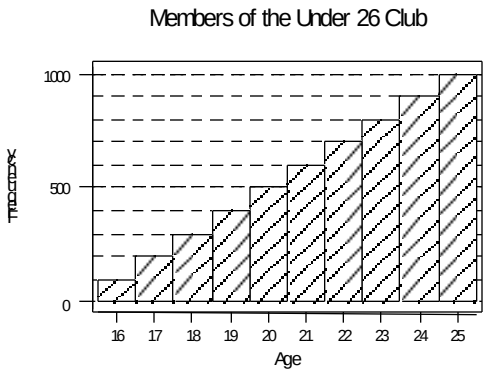
5. (15 points) The histogram on the right depicts the age of a sample of 5500 young Americans who are members of the "Under 26 Club". What **percent,** (not how many) of them were (2 point each):

_____ 21 or older.

_____ older than 17 but younger than 21.

_____ younger than 21.

_____ below 18 or above 22 combined.



_____ What is the level of this measurement for these Ages (2 points)?

Describe the **shape** of the distribution of these Ages in statistical terms (3 points).

_____

_____

_____

_____

Clearly identify the population (2 points).

_____

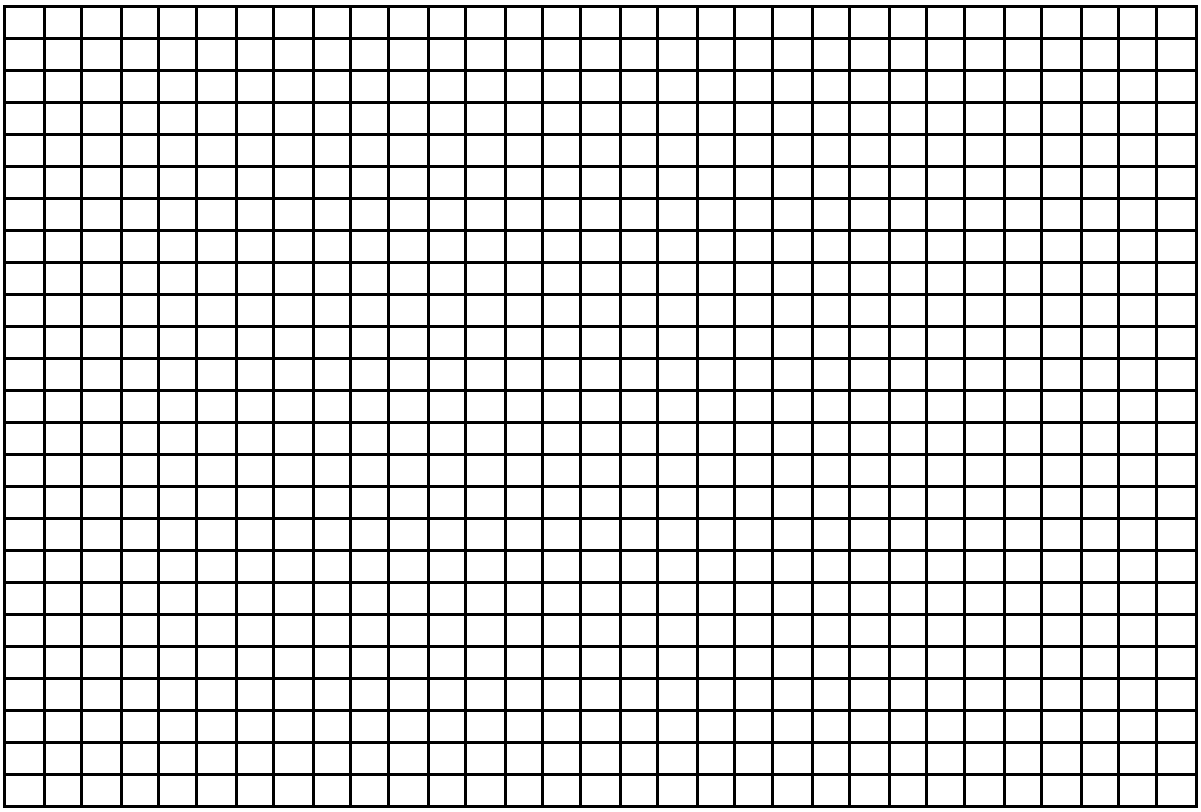
_____

_____

6. (41 points) The number of fatalities per 100,000 drivers and the median age of the driver are as follows:

Driver Median Age	17.5	22.5	29.5	39.5	49.5	59.5	67.0	79.5
Number of Fatalities	64.6	45.0	32.0	26.4	22.2	20.9	19.5	26.7

(8 points) On the grid below make a scatter plot of this data using median age on the *x*-axis. Label your axes. Make sure you use the entire grid.



Describe the overall pattern of the graph of this data and comment on whether or not you feel there are any outliers. Identifier all outliers by their coordinates (age, fatalities). (2 points)

_____

_____

_____

_____

_____

_____

_____ What is the explanatory variable for this data? (2 points)

_____ What is the level of measurement for the explanatory variable? (2 points)

_____ What is the response variable for this data? (2 points)

_____ What is the level of measurement for the response variable? (2 points)

_____ What is the value of the correlation coefficient? (4 points)

Briefly explain why the correlation should not be used here (4 points): _____

_____

_____

_____

_____

_____ Find the regression equation that best fits the scatter plot. Linear, Quadratic, or Exponential (9 points) (Round to 4 significant figures, i.e. 1407.525 rounds to 1408 and .005103465 rounds to .005103)

_____What percent of the variation is explained by the regression equation? (2 points)

_____Use your regression equation to estimate the number fatalities for a 20 year old driver. (2 points)

_____According to your regression equation, in approximately what is the median driving age(s) that had 30 fatalities per 100,000 drivers? Round off your answer to the **nearest tenth**. (2 points)

Name:_____ **Due:**_____ **Points Earned:**_____

I. (45 Points Total) How much wood can a woodchuck chuck, if a woodchuck could chuck wood? In the ancient land of Cecilville, all of the woodchucks can chop wood. Cecilia, the queen of Cecilville, wanted to know the answer to this question for her queendom, so she asked her faithful servant, the renowned statistician, Trapper John, to help her answer this question. To provide Cecilia with an answer, Trapper John placed woodchuck traps at 29 sites throughout the queendom, trapping a total of 1241 woodchucks. Each woodchuck trapped performed the task of chopping wood for as long as he or she was able, and the number of chords each woodchuck chopped was recorded. Trapper John discovered that the woodchucks selected chopped on average 1.52 chords of wood. Based on Trapper John's survey, Cecilia declared that Cecilville's woodchucks are capable of chopping approximately an average of 1-1/2 chords of wood apiece. This pleased her so much that she declared February 2 to be Woodchuck Day and made it an annual holiday. (Note: a chord of wood is defined to be a pile of wood that measures 4' x 4' x 8' and you can assume that it can be measured accurately.).

Directions: Answer questions 1 to 6 on a word processor numbering your responses accordingly. Save your word processor file in **Microsoft Word format.** Attach a printout of your responses to this sheet and email a copy of your **Microsoft Word** computer file to your instructor at jcliment@cecilcc.edu. Use "Trapper John-Your Name" as the subject of your email and label your computer file "Trapper John-You Name". Put your name and class on the top of your document. For example John Smith would label his file and email as: ***Trapper John – John Smith.***

1. (5 Points) In the space below, completely describe Trapper John's method of collecting data.

2. (5 Points) Identify the numerical descriptions (summaries) Trapper John gave to his data.

3. (5 Points) Completely and clearly describe the population from which Trapper John drew his sample.

4. (5 Points) What characteristic (describe the variable or the quantity) of the Population is being measured?

5. (5 Points) Completely and clearly describe the sample and tell its size.

6. (10 points) In a brief paragraph argue whether or not Trapper John's sample was a random sample. Carefully check the definition of a random sample on your course handout before answering this part.

7. (10 points) State whether the following statements are inferential or descriptive. **Write your answer in the blanks below.**

_____ Queen Cecilia claimed that, on average, the woodchucks in her queendom could chop approximately 1-1/2 chords of wood apiece.

_____ Trapper John found that the woodchucks trapped chopped on average 1.52 chords of wood apiece.

II. (20 Points Total) For each of the statements listed below, supply the proper statistical symbol for the quantity described in bold print. Essentially you are being asked to classify the numerical quantity given in **bold type** as a population proportion, population mean, sample proportion, sample mean, population size, sample size, or none of the above. If your answer is none of the above, place an "NA" in the space provided. Otherwise place the correct statistical symbol for the quantity described. The symbols can be found on your "Definitions and Formulas" handout. Place your answer in the blank on the left.

_____**1.** The State Police reported that **39%** of the fatal accidents in Cecil County last year were alcohol related.

_____**2.** The State Police reported that there were had **56** fatal accidents in Cecil County last year.

_____**3.** To determine future voting patterns, **998** Cecil County residents were polled about their voting plans.

_____**4.** The voting survey indicated that **62.4%** of the Cecil County residents polled said that they planned to vote.

_____**5.** The average grade on the first test of all twenty statistics students was **88.5**.

_____**6.** The average grade on the first test for the first five finishing students out of all twenty was **91.6**.

_____**7.** Donna's grade on the first test was **86**.

_____**8.** Based on a survey of 10,000 dentists, the average annual salary was reported to be **$76,224**.

_____**9.** The elections board reported that **75.4%** of the eligible voters in Cecil County were registered to vote.

_____**10.** The 2000 census showed that the average size of all households was **1.9** children per family.

III. (30 Points) The number of deaths from cancer in the United States has risen steadily over time. In 1991, for example, about 514,000 people died of cancer, up from 331,000 deaths in 1970. A member of Congress claims that these numbers show that "no progress has been made in treating cancer".

Directions: Answer questions 1 to 3 on a word processor numbering your responses. Save your word processor file in **Microsoft Word format.** Attach a printout of your responses to this sheet and email a copy of your **Microsoft Word** computer file to your instructor at jcliment@cecilcc.edu. Use "Cancer Deaths-Your Name" as the subject of your email and label your computer file "Trapper John-You Name". Put your name and class on the top of your document. For example John Smith would label his file and email as: ***Cancer Deaths – John Smith.***

1. Explain how it is possible for the number of people dying of cancer in the United States to increase over time, even while treatment of the disease is improving.
2. What variable, in an algebraic sense, did the congressman use to measure the effectiveness of cancer treatments?
3. What would have been the best or most appropriate variable for the congressman to use with this data in order to more accurately measure the effectiveness of medical treatments for this potentially fatal disease?

Name:_____ **Due:**_____ **Points Earned:**_____

IV. (30 Points.) According to the attached article, young adults weigh more today than they did in the past. Read the article and answer the following questions.

Directions: Answer questions 1 to 6 on a word processor numbering your responses accordingly. Save your word processor file in **Microsoft Word format.** Attach a printout of your responses to this sheet and email a copy of your **Microsoft Word** computer file to your instructor at jcliment@cecilcc.edu. Use "Adults Weight-Your Name" as the subject of your email and label your computer file "Trapper John-You Name". Put your name and class on the top of your document. For example John Smith would label his file and email as: ***Adults Weight – John Smith.***

1. (5 pts.) What is the level of measurement for the weight gain mentioned in the first paragraph?

2. (5 pts.) Is that measurement a parameter or a statistic?

3. (5 pts.) A second variable, race/gender, was used to further analyze this data. What is the level of measurement of this second variable?

4. (5 pts.) Dr. Lewis found that the average weight of young adults was 171 pounds. In one or two sentences, carefully and completely describe the population (in a statistical sense) to which Dr. Lewis is referring.

5. (5 pts.) In one or two sentences, carefully and completely describe the sample (including size, makeup, location, etc.) that was used to draw the conclusions in the article.

6. (5 pts.) The article implied that the weight gain for young adults was alarming. A student suspects that young adults were taller in '91-'92 than they were in '85-'86. If the student is correct, explain how this could affect this interpretation of the results.

Despite Better Diets, Young Adults Weighing More

TAMPA, Fla., March 17 (AP) — The average weight of young American adults rose 10 pounds in the last seven years despite a healthier diet, the Government reported today.

"This is bad news," said Dr. Cora E. Lewis, an epidemiologist at the University of Alabama in Birmingham who conducted the study.

"It's not just people getting older and heavier, but the population is getting heavier," Dr. Lewis told a meeting of the American Heart Association.

The director of the Federal study, Dr. Diane Bild of the National Institutes of Health, said the result "was totally unexpected."

Cause Not Determined

It was particularly surprising because the American diet has improved in recent years and cholesterol levels have been dropping.

In 1992-93, Dr. Lewis found that the average weight of adults from age 25 to age 30 was 171 pounds; in 1985-86, the average in that age group was 161 pounds.

Weight normally increases with age, but Dr. Lewis said the weight changes she had found were unrelated to aging. "That amount of change in that time period is shocking," she said.

The study did not address the cause of the weight gain, but Dr. Lewis speculated that a decline in physical activity was probably responsible.

A separate study, by the National Center for Health Statistics, has found that Americans are consuming more calories, though the amount of cholesterol and fat in the American diet is dropping, according to a recent summary in The Morbidity and Mortality Weekly report, a publication of the Centers for Disease Control and Prevention in Atlanta.

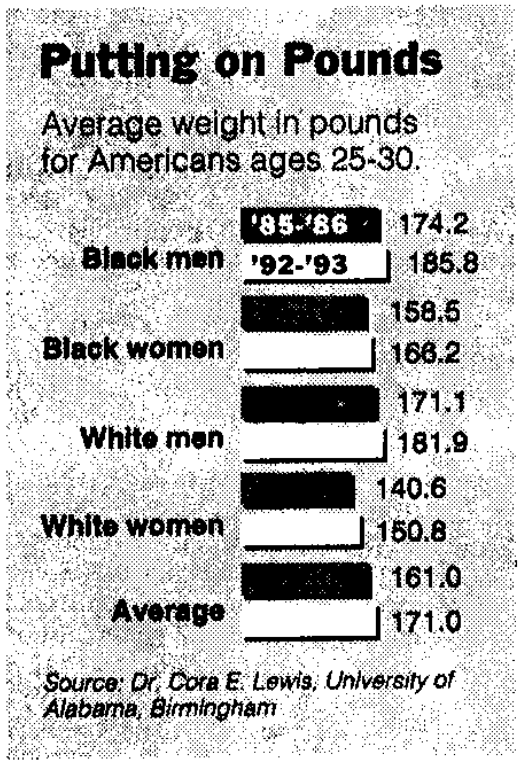
Advice to the Young

Becky Huff Lankenau, a nutritionist at the centers, said the new study illustrated the need for educating and encouraging young people to exercise.

"Probably the safest and most positive thing would be to nurture the idea of physical activity," she said.

The study by Dr. Lewis and Dr. Bild and their colleagues was based on data on 5,115 people, equally divided among blacks and whites and men and women. The participants came from Chicago, Minneapolis, Birmingham and Oakland, Calif.

While weight climbed in the group in 1992-93, cholesterol levels dropped 9 points, probably because the diets had less cholesterol and saturated fat, Dr. Bild said. She added that such diet improvements should reduce the heart-disease risks of those young adults but that it was not known whether the weight gain would counteract that benefit.



Name:_____ **Due:**_____ **Points Earned:**_____

You will need the Class Minitab Worksheet 'Pennies–Fall2002'.

V. (10 Points Total) In the space below, draw a stem & leaf plot of the weights using the hundredths digit as the leaf and a stem interval of length 0.02 (2 hundredths). Note: you can do this on Minitab and copy it to this sheet, but make sure you specify the interval increment correctly.

Stem | Leaves

VI. (10 points) In the space below, based on your stem and leaf plot describe the most important feature observe about your dataset**.** **PRINT YOUR ANSWER**.

VII. (10 Points) On separate paper, draw a **bar type** time-series plot (not a line type) of the frequency of the pennies in our sample. List each year individually not groups of years.

VIII. (40 Points Total) We often look at time series data to see the effect of a social change or new policy. Here are data on motor vehicle deaths in the United States. Because motor vehicle deaths will tend to rise as motorists drive more miles, we look instead at the rate of deaths, which is the number of deaths per 100 million miles driven.

Year	Rate	Year	Rate	Year	Rate	Year	Rate
1960	5.1	1968	5.2	1976	3.3	1984	2.6
1962	5.1	1970	4.7	1978	3.3	1986	2.5
1964	5.4	1972	4.3	1980	3.3	1988	2.4
1966	5.5	1974	3.5	1982	2.8	1990	2.2

_____ **1.** (5 points) What is the level of measurement for Rate?

_____ **2.** (5 points) What is the level of measurement for Year?

_____ **3.** (5 points) Is this study observational or experimental?

4. On separate paper make a **line-type** time series plot (not a bar chart type) of this death rate data and describe the overall pattern of this data (how it varies over time). Use year as the independent variable and label your axes. Make sure you maximize the use of your y-axis scale when you draw your graph. Part of the purpose of this exercise, is to see if you can draw graphs by hand, so only hand drawn graphs will be accepted.

Directions: Answer questions 5 to 7 on a word processor numbering your responses accordingly. Save your word processor file in **Microsoft Word format.** Attach a printout of your responses to this sheet and email a copy of your **Microsoft Word** computer file to your instructor at jcliment@cecilcc.edu. Use "Traffic Deaths-Your Name" as the subject of your email and label your computer file "Trapper John-You Name". Put your name and class on the top of your document. For example John Smith would label his file and email as: ***Traffic Deaths – John Smith.***

5. (5 points) Give a brief verbal description of overall pattern.

6. (5 points) In 1974 the national speed limit was lowered to 55 miles per hour in an attempt to conserve gasoline after the 1973 Mid-east war. In the mid-1980s most states raised speed limits on interstate highways to 65 miles per hour. Some said that the lower speed limit saved lives. Explain if the effects of the lower speed limits between 1974 and the mid-1980s are visible in your plot.

7. (5 points) Suppose that the raising of the speed limit in the mid-1980s actually caused more deaths than the lower speed limit. Furthermore, suppose that deaths due to accidents continued to drop in the decade following the higher speed limit. Give an explanation of how this could happen.

Name:_____ **Due:**_____ **Points Earned:**_____

You will need the following datasets: 'Pennies-Fall2002' & 'Peanuts-Fall2002'.

IX. (45 points) For the **sample** of Pennies given in the dataset 'Pennies-Fall 2002' find the following quantities for their weight. Note: If necessary, round off your **final calculation (not your intermediate calculations)** to the **nearest hundredth.**

_____ Find Σx.

_____ Find Σx^2.

_____ Calculate the mean weight.

_____ Calculate the median weight.

_____ Calculate the 10% trimmed mean weight.

Note:
Round Final Calculation to Nearest Hundredth

_____ Calculate the variance of the weight by the computational formula.

_____ Calculate the standard deviation of the weight by the computational formula.

_____ Calculate the mean absolute deviation of the weight.

_____ Calculate Q_3 using your **T**ext's definition, **M**initab's rule or your **C**alculator. Since these may be different place a capital **T**, **M**, or **C** next to your answer to indicate which rule you used.

_____ Calculate Q_1 by the method used above.

_____ Calculate the interquartile range by the method used above.

_____ List all outliers using the text definition on page 108. Write 'none', if there are no outliers.

_____ Calculate the coefficient of variation of the weight.

_____ Calculate P_{10} for the weight.

_____ Calculate P_{33} for the weight

X. (10 points) In the space below, carefully draw a **sideways** box-plot of the weights of pennies. Label your axis.

XI. (35 Points Total) For the '**Peanuts-Fall2002'** dataset find descriptive statistics for both the **one-nut** and **two-nut** type peanuts.

1. (20 points) In the space below, draw back-to-back stem and leaf graphs of these two data sets. Round your data to the nearest tenth and use a leaf unit of 0.1 with two leaves per grouping.

Leaves for One-Nut Type	Stem	Leaves for Two-Nut Type

2. (5 points) It should be clear from your stem and leaf plots that the one-nut peanuts weigh less than the two-nut peanuts, and that the two groups differ in size and perhaps shape. Another difference in the characteristics of these two data sets should be obvious. In one sentence identify the characteristic and state how the data sets differ in this characteristic.
PRINT YOUR ANSWER LEGIBLY BELOW.

3. (5 points) Someone claimed that typical two-nut peanuts should weigh twice as much as the one-nut peanuts. Use the descriptive statistics you found to explain whether or not you think this claim is correct. Make sure you identify the statistics you use in support of your answer. **PRINT YOUR ANSWER LEGIBLY BELOW.**

4. (5 points) Someone claimed that the variability of typical two-nut peanuts should be twice that of one-nut peanuts. Use the descriptive statistics you found to explain whether or not you think this claim is correct. Make sure you identify the statistics you use to support your answer. **PRINT YOUR ANSWER LEGIBLY BELOW.**

Name:_____ **Due:**_____ **Points Earned:**_____

X1I. (10 points) A large painting company hired Randy as a painter. They first gave him a test to see whether it would be best for the company to assign him to paint walls, trim, ceilings, or floors. The test consisted of her performing each of these three types of painting and measuring how long it took him to complete the task. Randy's results along with those of all previous candidates are summarized in the following table. In the first column on the table calculate the z-score for Randy's performance.

Randy's z-score	Randy's Results	Type of Painting	Mean of all Previous Applicants	Standard Deviation of all Previous Applicants
	51.8 min.	**Walls**	**39.06 min.**	**5.2 min.**
	1.7 hrs..	**Trim**	**5.55 hrs.**	**2.2 hrs.**
	63 min.	**Ceilings**	**22.50 min.**	**18.0 min.**
	98.1 min.	**Floors**	**109.332 min.**	**7.2 min.**

_____ Based on Randy's z-scores, in which type of painting should he be assigned to do?

XIII. (37 Points Total) The following problems are based on the Adventure Learning Systems Problems.

1. (10 points) A sample of 143 high school students was asked to give their weekly salary. Eighty of the students did not work so they responded that their weekly salary was zero. For the remaining students, their average weekly reported salary was \$48 with \$153 being the largest weekly salary reported. Answer the following questions and show your work in the space provided. Write **"MORE INFO"** for your answer, if there is not enough information to determine the answer. **Round off all answers to the nearest hundredth.**

_____ What is the range of this data?

_____ What is the mode of this data?

Round to Nearest Hundredth.

_____ What is the median for this data?

_____ What is the average of all of the respondents?

_____ What proportion of the respondents worked?

2. (12 points) Sally made 92 phone calls on her cellular phone last month. The average cost of her calls was \$4.00 with a standard deviation of \$1.60. Her most expensive phone call was \$8.04. Fourteen percent of her calls were less than \$2.06 and the third quartile, Q_3, was \$5.12. Answer the following questions and show your work in the space provided. Write **"MORE INFO"** for your answer, if there is not enough information to determine the answer. **Round off all answers to the nearest hundredth.**

_____ What is the z-score for the phone call that costs \$2.48?

_____ What is the price of the phone call that had a z-score of –1.05?

_____ What is the range for this data? **Round to nearest hundredth.**

_____ What is the median for this data?

_____ To the nearest whole number, calculate the approximate number of calls that cost at least \$2.06.

_____ To the nearest whole number, calculate the approximate number of calls that cost less than \$5.12.

3. (15 points) A bank has 244 customers with balances from $0 to over $150,000. Two of their customers are listed in the table below.

Customer	Balance	Percentile	z-score
Bill	$11,500	24	-1.50
Janice	$80,000	55	0.45

_____ Which customer has a balance that is closest to the mean?

_____ To the nearest whole number, calculate the approximate number of customers with balance below $80,000.

_____ To the nearest whole number, calculate the standard deviation for this data.

XIV. (20 points) The semester averages of a sample of 40 of the students in a fictitious class of statistics students are summarized in the following relative frequency table. Find the mean, variance, and standard deviation for the semester averages for the students in this class. Note: you will first have to calculate the frequencies and the midpoints for each interval. Several blank columns have been included for any additional calculations you may wish to make. Note: If necessary, round off your **final calculation (not your intermediate calculations)** to the **nearest hundredth.**

Interval	Midpoint	Relative Frequency	Frequency			
63.5 - 72.5		**0.10**				
72.5 - 81.5		**0.15**				
81.5 - 90.5		**0.45**				
90.5 - 99.5		**0.30**				
Total		**1.00**	**40**			

_____ Mean

_____ Variance

Round final answers to the nearest hundredth.

_____ Standard Deviation

_____ What is the level of measurement for this data?

XV. (40 points total) Revisit the Pennies data set.
1. (20 points) Using year as your independent variable, use Minitab to draw a plot of weight versus year. Attach your plot to this sheet.

2. (20 points) Recall that the stem-and-leaf diagram found previously for this data showed an important feature. From your Minitab plot you should now be able to observe the cause of this feature. In the space below, give an explanation of the cause.

Name:_____ **Due:**_____ **Points Earned:**_____

XVI. In February 2000, The New York Times published data on Nokia sales. For the years 1991 through 1997, the sales in billions of dollars are shown in the table below.

Sales	3.8	3.6	4.0	6.2	8.2	8.2	9.7
Year	91	92	93	94	95	96	97

1. (85 points) Answer the following questions.

_____What is the correlation coefficient between Sales and Year?

_____What is the equation **(use three significant figures in rounding your coefficients)** of the linear regression line for the sales, using year as the predictor? **Use the data in the form listed.** **(If you round off this equation incorrectly or if you get the explanatory and response variables wrong, many of your answers may be judged incorrect.)**

_____What Variable did you use for your explanatory variable?

_____Did you understand the meaning of three significant digits? (If not, ask for help!)

_____What is the value of R^2 for the above regression equation?

_____What percent of the variation is explained by the above regression equation?

_____Based on the **rounded-off** regression equation, what were Nokia's sales **in billions of dollars** for 1995?

_____Based on the **rounded-off** regression equation, what were Nokia's sales **in billions of dollars** for 1980?

_____Based on the **rounded-off** regression equation, what will Nokia's sales be **in billions of dollars** for 2000? **(Note: what value should you use for the year 2000 to make sure that it numerically follows 99?)**

_____The actual Nokia sales for 2000 were 27.0 billions of dollars. Find the difference between this actual value and your predicted value above.

Based on the difference in the actual and predicted Nokia sales found above, explain what this tells you about the use of a linear model here. **Print your answer neatly below.**

_____

_____

_____

_____

_____

_____Based on the **rounded-off** regression equation, in what year did Nokia sales first reach 1 billion dollars?

_____What is the mean sales in billions of dollars for the years listed? **Round to three significant figures.**

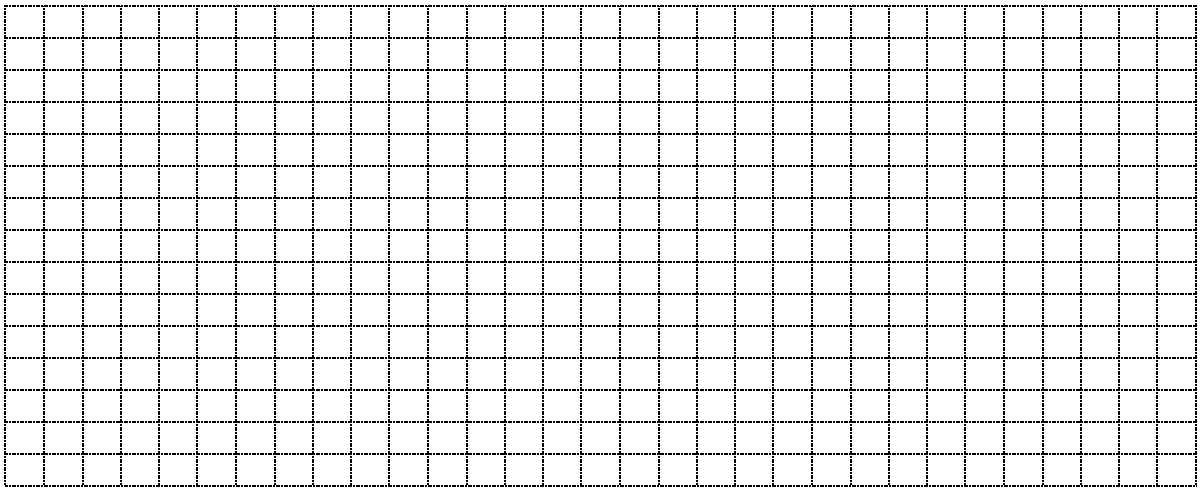
_____Based on the **rounded**-off regression equation, **rounded-off mean**, calculate the year in which Nokia sales first reached the mean sales.

_____What is the median sales in billions of dollars for the years listed? **Round to three significant figures.**

_____Based on the **rounded**-off regression equation, **rounded-off median**, calculate the year in which Nokia sales first reached the median sales.

From a **statistical point of view**, explain why the predicted Nokia sales you calculated for 1980 could be misleading. **Print your answer neatly below.**

2. (35 points) According to The New York Times Nokia sales for the years 1998 through 2001 were respectively in billions of dollars 15.4, 20.0, 27.0 and 28.0. Add these sales to your dataset and on the grid below, draw a scatter plot of this revised data set. **(Note: using the convention given, 98 for 1998, 99 for 1999, etc., what values should you use for the years 2000 and 2001 to ensure that the numbers used are greater in sequence than 98 & 99?)**



_____Based on your scatter plot and the expected behavior of these two variables, tell what type of regression equation, linear, quadratic, cubic, or exponential, should be used to best fit this data?

Note: if Model is Incorrect your remaining answers will be judged incorrect.

_____ Use Minitab, or your Calculator to find the regression equation of your curve. **Round all of your coefficients to three significant figures.**

_____What is the value of R^2 for the above regression equation?

Note: if you rounded off incorrectly, your remaining answers may be judged incorrect.

_____Based on the **rounded-off** regression equation, what are the expected Nokia sales in billions of dollars for 2000?

_____Find the difference between this actual Nokia sales for 2000 and your predicted value above.

_____Based on the **rounded-off** regression equation, in what year did Nokia sales first reach 25 billion dollars?

I. 1. Trapper John placed traps at 29 sites throughout Cecilville and trapped 1241 woodchucks in them. The woodchucks "selected" performed the task of chopping wood as long as they were able and the number of chords or wood that each one chopped was recorded.

2. The average number of chords of wood chopped by the 1241 woodchucks "selected" was 1.52 chords.

3. All the woodchucks of Cecilville.

4. The amount of wood, measured in chords, that a woodchuck is capable of chopping at one time.

5. 1241 woodchucks were "selected" from 29 sites throughout Cecilville.

6. This is not a random sample. For a sample to be a random sample, each subject or unit in the population must have an equally likely chance of being selected. For example, if the population consists of 500 individuals, then each individual must have a 1/500 (0.002 or 0.2%) chance of being selected. It is nearly impossible for all of the woodchucks of Cecilville had an equally likely chance of being caught in the traps. Woodchucks that lived in the vicinity of the 75 sites were more likely to be trapped than those from other areas were. Those that never left their homes could not be trapped unless the traps were placed in their homes. There are numerous other reasons why some woodchucks had a greater likelihood of being caught than others. How many can you name?

7. Inferential
Descriptive

II.

1. p
2. N
3. n
4. $\hat{p}$
5. μ
6. $\bar{x}$
7. NA
8. $\bar{x}$
9. p
10. μ

III. 1. There are several reasons why the number of deaths due to cancer could be increasing while treatment of the disease is improving.

- Cancer is a disease that proportional afflicts older people more than younger people. Since the number and proportion of older people in our population has increased over the years, it is very reasonable for the number of deaths to increase even as treatment improves. In addition, since the total population is increasing, cancer deaths could increase with improved treatment.
- The population of the U.S. increased from 1970 to 1991, therefore the number of cancer patients would increase along with the population.
- Pollution or other environmental factors may be increasing the incidence of cancer, and consequently the number of deaths, even if treatment is improving.

2. The variable that the congressman used was the **number of people dying of cancer.**

3. Possible answers.

- First of all, one should to look at each age group individually. For each age group, one could then measure the **ratio or percentage of people afflicted with cancer who die (or survive) from the disease**. This variable should give us a clearest picture as to whether or not the treatment of cancer is improving.
- One could also look at the **overall population death rate due to cancer (or the survival rate)** for each age group; however, changes in this variable might be due to changes in a different variable: **the incidence of cancer**. In the language of statistics, if the cancer death rate per person in the population were used to measure the effectiveness of treatment, then we would say that **"improved treatment"** was **confounded** with **"incidence of cancer"**. Here we would not be able to distinguish which of these two variables was the cause of any change observed in the cancer death rate.
- Another variable one could look at is the **average life span after diagnoses**; however, it is very possible that in 1991 doctors were able to diagnose patients' cancers much earlier they could in 1970. Thus, an increase in the life span after diagnoses might be due to earlier diagnoses and not due to improved treatment.

What is most alarming about this problem is that it was a real live U.S. Congressman who made this statement. Hopefully, he was merely imprecise in his use of the English language. Perhaps he spoke of the "number" of people dying of cancer, when he really meant the "percentage" of people dying of cancer. One would hope that someone elected to such a lofty position would know the difference between these two concepts. If the congressman did not understand that the cancer rate was the appropriate measure to use, then the electorate should not return him to office. Perhaps he could find a job caring for cancer patients.

IV. 1. Ratio
2. Statistic
3. Nominal

4. There are three key items (in bold type) that identify the population. The population consists of the weights of **young Americans** between the age of **25 & 30** in the years **1992-1993.**

5. There are four key items (in bold type) that identify sample. The sample consists of **5,115 people**, between the ages of **25-30**, equally divided among **blacks and whites, and men and women**. The participants came from **Chicago, Minneapolis, Birmingham and Oakland.**

6. The article implied that the weight gain for young adults was alarming. If people were taller 1n 1992-1993, than they were in 1985-86, then the weight gain observed could completely or partially be attributed to this gain in height. As a result, the interpretation that the weight gain is alarming might no longer be valid. If weight gain was simply be due in whole or part to a gain in stature then we would not be as alarmed by it.

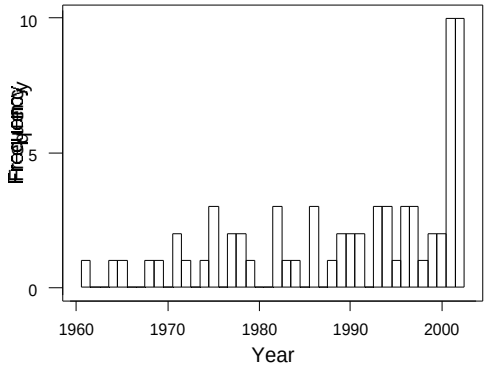
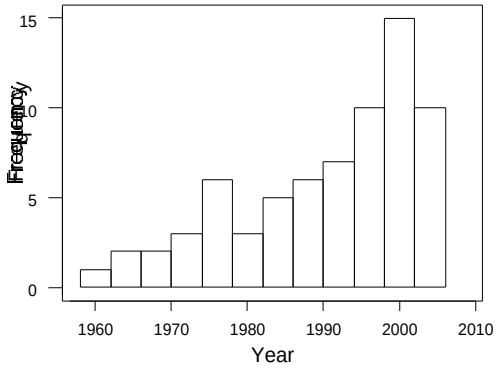
V. (10 Points Total) In the space below, draw a stem & leaf plot of the weights using the hundredths digit as the leaf and a stem interval of length 0.05 (5 hundredths).

2 Hundredths	5 Hundredths
Stem-and-Leaf Display: Weight Stem-and-leaf of Weight N = 70 Leaf Unit = 0.010 3 24 667 24 24 888888899999999999999 (16) 25 0000000001111111 30 25 222223 24 25 45 22 25 6 21 25 9 20 26 20 26 20 26 20 26 20 26 20 27 20 27 20 27 20 27 20 27 20 28 20 28 20 28 20 28 20 28 20 29 20 29 20 29 20 29 20 29 20 30 20 30 20 30 45 18 30 666777 12 30 8899 8 31 00111 3 31 2 2 31 44	**Stem-and-Leaf Display: Weight** Stem-and-leaf of Weight N = 70 Leaf Unit = 0.010 24 24 667888888899999999999999 (23) 25 00000000011111112222234 23 25 569 20 26 20 26 20 27 20 27 20 28 20 28 20 29 20 29 20 30 4 19 30 56667778899 8 31 00111244

VI. (10 points) In the space below, based on your stem and leaf plot describe the most important feature observe about your dataset**.**

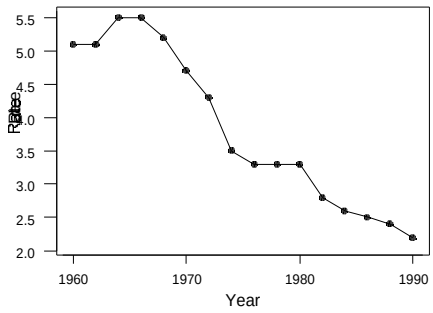
There appears to be two groups of pennies, those that weigh a little over 3 grams and those that weigh around 2.5 grams. This is a classical bimodal distribution, which indicates that we have two separate populations for the weights of the pennies.

VII. (10 Points) In the space below, draw a **bar type** time-series plot (not a line type) of the frequency of the pennies in our sample for each year individually not groups of years.

Minitab with Controlled Intervals of 1 Year	Minitab Default
	

VIII. 1. Rate is **Ratio.** **2.** Year is **Interval**. **3. Observationa**l.

4.



5. The overall pattern:

The accident rate rises slightly during the early to mid 1960's, reaching a peak in 1968. It then falls off steeply and steadily until 1976 where it levels off until 1980. During the 1980's the rate resumes its decline, steadily but not as steeply as the earlier decline.

6. Is Effect of Lower Speed Limit Visible?

No it is not visible. The accident rate was declining steeply prior to 1974 when speed limits were reduced to 55 mph. One would expect the effect of this reduced speed limit would be to further reduce the accident rate; however, just the opposite is seen. The accident rate levels off for most of the period during which the new speed limit was in effect. Although we were not asked to comment on the raising of speed limits in the middle 1980's, its effect is also not evident. Logic tells us that the accident rate should increase with higher speed limits, but the graph shows a decrease.

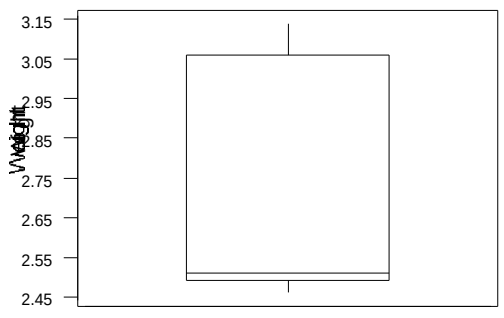
7. Although the higher speed limit may be causing more vehicle accident deaths, there could be several variables that may be having the opposite effect. Here are some possibilities:

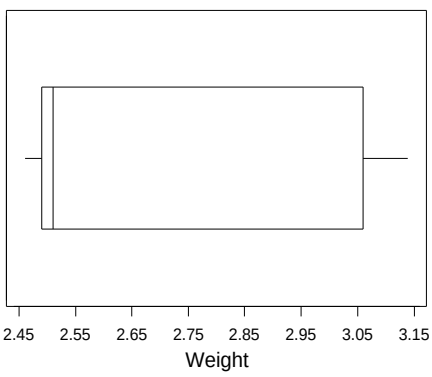
- Vehicles may be safer in a crash, due to better design or the addition of safety devices such as air bags.
- More people may be using safety devices, such as seat belts and car seats for children.
- Perhaps there is better enforcement of the traffic laws.
- Roads themselves might be safer.
- People may be driving fewer miles than they did before.
- Anti drinking and driving campaigns may have reduced the number of impaired drivers on the road.

IX. (45 points) For the **sample** of Pennies data set, for the weight of the pennies, find the following quantities. Note: If necessary, round off your **final calculation (not your intermediate calculations)** to the **nearest hundredth.**

Text	Minitab	TI-83	
186.80	186.80	186.80	Find Σx.
503.45	503.45	503.45	Find Σx^2. **503.4462**
2.67	2.67	2.67	Calculate the mean weight. **2.668571429**
2.51	2.51	2.51	Calculate the median weight.
2.64	2.65	2.64	Calculate the 10% trimmed mean weight. **Round Final Calculation to Nearest Hundredth**
0.07	0.07	0.07	Calculate the variance of the weight by the computational formula. **0.0708151021**
0.27	0.27	0.27	Calculate the standard deviation of the weight by the computational formula. **0.2661110709**
0.24	0.24	0.24	Calculate the mean absolute deviation of the weight. **Using rounder or unrounded mean.**
3.06	3.06	3.06	Calculate Q_3 using your **T**ext's definition, **M**initab's rule or your **C**alculator. Since these may be different place a capital **T**, **M**, or **C** next to your answer to indicate which rule you used.
2.49	2.49	2.49	Calculate Q_1 by the method used above.
0.57	0.57	0.57	Calculate the interquartile range by the method used above.
None	None	None	List all outliers using the text definition on page 108. Write 'none', if there are no outliers.
9.97 or 10.11	9.97 or 10.11	9.97 or 10.11	Calculate the coefficient of variation of the weight. **.2661110709/2.66857149*100=9.972042042 or .27/2.67*100=10.11**
2.48	2.48	2.48	Calculate P_{10} for the weight. **Average of 7th & 8th Data Point**
2.49	2.49	2.49	Calculate P_{33} for the weight **24th Data Point**

X. (10 points) In the space below, carefully draw a box-plot of the weights of pennies. Label your axis.





XI. (35 Points Total) For the **Peanuts** dataset find descriptive statistics for both the **one-nut** and **two-nut** type peanuts.

1. (20 points) In the space below, draw back-to-back stem and leaf graphs of these two data sets. Round your data to the nearest tenth and use a leaf unit of 0.1 with two leaves per grouping.

Leaves for One-Nut Type	Stem	Leaves for Two-Nut Type
1 1 1	1a	
3 2 2 2 2 2 2	1b	
5 5 5 4 4 4 4 4 4	1c	4
7 7 7 7 7 7 6 6 6 6 6 6	1d	6
9 9	1e	8 9
	2a	0 1 1
	2b	2 2 2 3 3 3 3
	2c	4 4 5 5 5 5
	2d	6 6 6 6 7 7 7
	2e	8 8 8 8 9
	3a	0 1
Done automatically in Minitab without first rounding the data.	3b	2 3
	3c	4
	3d	7

Leaves for One-Nut Type	Stem	Leaves for Two-Nut Type
1 1	1a	
3 3 3 3 3 3 2 2	1b	
5 5 5 5 5 5 4 4	1c	5
7 7 7 7 7 7 6 6 6 6	1d	7
9 9 8 8	1e	8 9
0	2a	1
	2b	2 2 2 3 3 3 3 3 3
	2c	4 5 5
	2d	6 6 6 6 7 7 7 7 7
	2e	8 8 9 9 9
	3a	0 0
Rounded first and then plotted.	3b	2 3
	3c	4 5
	3d	7

2. (5 points) It should be clear from your stem and leaf plots that the one-nut peanuts weigh less than the two-nut peanuts, and that the two groups differ in size and perhaps shape. Another difference in the characteristics of these two data sets should be obvious. In one sentence identify the characteristic and state how the data sets differ in this characteristic.
PRINT YOUR ANSWER BELOW.

There is more variability in the two-nut peanuts than in the one-nut peanuts.

3. (5 points) Someone claimed that typical two-nut peanuts should weigh twice as much as the one-nut peanuts. Use the descriptive statistics you found to explain whether or not you think this claim is correct. Make sure you identify the statistics you use to support your answer. **PRINT YOUR ANSWER BELOW.**

Descriptive Statistics: Weight by Number

Variable	Number	N	Mean	Median	TrMean	StDev
Weight	1	33	1.5097	1.5100	1.5055	0.2314
	2	38	2.5724	2.5800	2.5712	0.4805

While Minitab's descriptive statistics (see above) shows that the mean and the median weight of the two-nut peanuts is considerable grater than that for he one-nut peanuts, neither of those numbers for the two-nut peanuts is twice as large as the corresponding number for the one-nut peanuts.

4. (5 points) Someone claimed that the variability of typical two-nut peanuts should be twice that of one-nut peanuts. Use the descriptive statistics you found to explain whether or not you think this claim is correct. Make sure you identify the statistics you use to support your answer. **PRINT YOUR ANSWER BELOW.**

Based on the standard deviations, Minitab's descriptive statistics shows that the two-nut peanuts are slightly more than twice as variable as the one-nut peanuts (0.4805 vs. 0.2314).

MAT 127 – Climent Test 1 – Take Home Answers Pages 9 – 10

X1I. (10 points) A large painting company hired Randy as a painter. They first gave him a test to see whether it would be best for the company to assign him to paint walls, trim, ceilings, or floors. The test consisted of her performing each of these three types of painting and measuring how long it took him to complete the task. Randy's results along with those of all previous candidates are summarized in the following table. In the first column on the table calculate the z-score for Randy's performance.

z-score		
2.45	Walls	
-1.75	Trim	**Best**
2.25	Ceilings	
-1.56	Floors	

XIII. (37 Points Total) The following problems are based on the Adventure Learning Systems Problems.

1. (10 points) A sample of 143 high school students was asked to give their weekly salary. Eighty of the students did not work so they responded that their weekly salary was zero. For the remaining students, their average weekly reported salary was $48 with $153 being the largest weekly salary reported. Answer the following questions and show your work in the space provided. Write **"MORE INFO"** for your answer, if there is not enough information to determine the answer. **Round off all answers to the nearest hundredth.**

153 What is the range of this data?
0 What is the mode of this data?
0 What is the median for this data?
21.15 What is the average of all of the respondents?
63/143 = 0.44 (0.44056) What proportion of the respondents worked?

2. (12 points) Sally made 92 phone calls on her cellular phone last month. The average cost of her calls was $4.00 with a standard deviation of $1.60. Her most expensive phone call was $8.04. Fourteen percent of her calls were less than $2.06 and the third quartile, Q_3, was $5.12. Answer the following questions and show your work in the space provided. Write **"MORE INFO"** for your answer, if there is not enough information to determine the answer. **Round off all answers to the nearest hundredth.**

– 0.95 What is the z-score for the phone call that costs $2.48?
-1.05(1.60)+4.00 = 2.32 What is the price of the phone call that had a z-score of -1.05?
More Info What is the range for this data?
More Info What is the median for this data?
.96(92)=88.32=88 To the nearest whole number, calculate the approximate number of calls that cost at least $2.06.
.75(92)=69 To the nearest whole number, calculate the approximate number of calls that cost less than $5.12.

3. (15 points) A bank has 244 customers with balances from $0 to over $150,000. Two of their customers are listed in the table below.

Customer	Balance	Percentile	z-score
Bill	$11,500	24	-1.50
Janice	$80,000	55	0.45

One needs to solve the following equations simultaneously for σ: _____ _____

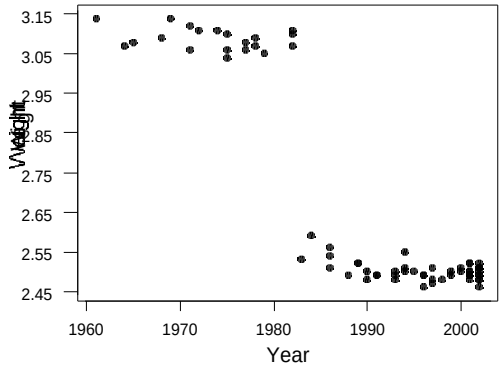
Janice Which customer has a balance that is closest to the mean?
134 (244 * 0.55 = 134.2) To the nearest whole number, calculate the approximate number of customers with balance below $80,000.
35138 (35128.2) To the nearest whole number, calculate the standard deviation for this data.

XIV. (20 points) The semester averages of a sample of 40 of the students in a fictitious class of statistics students are summarized in the following relative frequency table. Find the mean, variance, and standard deviation for the semester averages for the students in this class. Note: you will first have to calculate the frequencies and the midpoints for each interval. Several blank columns have been included for any additional calculations you may wish to make. Note: If necessary, round off your **final calculation (not your intermediate calculations)** to the **nearest hundredth**

Interval	Midpoint	Rel. Frequency	Frequency	M*F	M^2*F	
63.5 - 72.5	**68**	0.10	**4**	**272**	**18496**	
72.5 - 81.5	**77**	0.15	**6**	**462**	**35574**	
81.5 - 90.5	**86**	0.45	**18**	**1548**	**133128**	
90.5 - 99.5	**95**	0.30	**12**	**1140**	**108300**	
Total		1.00	40	**3422**	**295498**	

3422/40 = **85.55** Mean
$(40*295498-3422^2)/(40*39)$ = 70.40769231= **70.41** Variance **Round final answers to the nearest hundredth.**
SQRT(70.40769231) = 8.390929168 = **8.39** Standard Deviation
Interval or Ratio What is the level of measurement for this data?

XV. (40 points total) Revisit the Pennies data set.
1. (20 points) Using year as your independent variable, use Minitab to draw a plot of weight versus year. Attach your plot to this sheet.



2. (20 points) Recall that the stem-and-leaf diagram found previously for this data showed an important feature. From your Minitab plot you should now be able to observe the cause of this feature. In the space below, give an explanation of the cause.

The weight of the pennies changed in 1983. Pennies made before 1983 weighed more than those made from 1983 on.

Problem 1

Descriptive Statistics

Variable	N	Mean	Median	TrMean	StDev	SE Mean	Minimum	Maximum	Q1	Q3
Sales	7	6.243	6.200	6.243	2.503	0.946	3.600	9.700	3.800	8.200

Correlations: Pearson correlation of Year and Sales = 0.959

Regression Analysis: The regression equation is:

Predictor	Coef	SE Coef	T	P
Constant	-98.16	13.89	-7.07	0.001
Year	1.1107	0.1477	7.52	0.001

S = 0.7815 R-Sq = 91.9% R-Sq(adj) = 90.3%

The regression equation is: **Sales = –98.2 + 1.11 Year**

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	34.543	34.543	56.56	0.001
Residual Error	5	3.054	0.611		
Total	6	37.597			

0.959 is the correlation coefficient between Sales and Year.

Sales = –98.2 + 1.11 Year is the correctly rounded equation of the regression line.

Year is the explanatory variable.

Yes, three significant digits means the first three digits beginning at the first non-zero digit.

0.919 or 91.9% is the value of R^2 for the above regression equation.

91.9% is the percent of the variation is explained by the above regression equation.

Sales = –98.2 + 1.11(95) = 7.25 is the expected sales for 1995.

Sales = –98.2 + 1.11(80) = –8.29 is the expected sales for 1980.

Sales = –98.2 + 1.11(100) = 12.8 is the expected sales for 2000.

27.0 – 12.8 = 14.2 is the difference in the actual and predicted sales for 2000.

This difference is too large to reliably use the linear model to extrapolate the results outside the data range (1991 to 1997) to the year 2000. Perhaps a different model is called for if extrapolation is our goal.

1990 The regression equation shows that sales for 1989 were 0.59 and in 1990 they were 1.70.

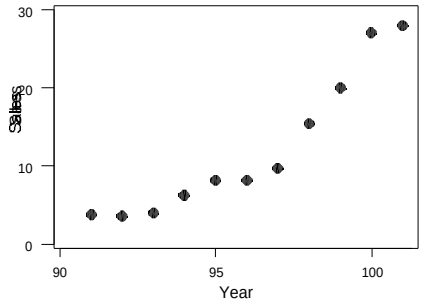
6.24 is the mean sales for the years listed.

1995 Year = (Sales + 98.2)/(1.11) = (6.24 + 98.2)/(1.11) = 94.09 (Round up).

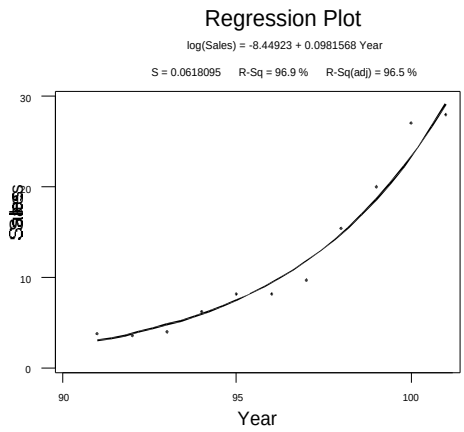
6.20 is the median sales for the years listed.

1995 Year = (Sales + 98.2)/(1.11) = (6.20 + 98.2)/(1.11) = 94.05 (Round up).

From a **statistical point of view** the predicted sales could be misleading, because it was calculated from the year 1980 that was not within the range of years (1991 to 1997) from which the regression equation was derived. Whenever one extrapolates outside the range of values from which the regression equation was derived, one runs the risk of getting erroneous information.



Problem 2



The regression equation is log(Sales) = -8.44923 + 0.0981568 Year

S = 0.0618095 R-Sq = 96.9 % R-Sq(adj) = 96.5 %

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	1.05982	1.05982	277.410	0.000
Error	9	0.03438	0.00382		
Total	10	1.09421			

Exponential is the equation that will best fit this data.

$\log_{10}(Y) = -8.45 + 0.0982\ X$ is the regression equation **TI-83 Answer: $y = 0.00000000355*(1.25)^X$.**

0.969 or 96.9% is the value of R^2 for the above regression equation.

23.4 is what you should expect sales for 2000. **TI-83 Answer: 17.4**

3.6 is the difference. **TI-83 Answer: 9.6**

2001 is when the predicted sales first reached 25 billion. **TI-83 Answer: 2002**

Page 1 Test 2 MAT 127 Climent Name: _____

Directions: 1. Show all work on these sheets. **Grade: _____**

2. **Calculators are permitted.**
3. **Computers are permitted and Minitab handouts may be used.**
4. **You may use hand written notes on one sheet of 8.5" by 11" paper.**
5. **You must hand in all written materials including your notes and scrap paper.**
6. **You must sign in and out for this test.**
7. **You may use the tables reproduced from your book that are available from the Math Lab.**
8. **You may take this test one page at a time.**

Show Your Work for any Partial Credit

1. (10 points) Suppose that a math test's grades are normally distributed. If the mean on the test was 77.8 with a standard deviation of 10, find the probability that a randomly selected student's grade on the test is: **SHOW YOUR WORK**

_____ at least 82.

_____ between 77 and 82.

2. (10 points) Suppose that in the previous problem, we select a random sample of 25 students, find the probability that their average grade on the test is:

_____ at least 82.

_____ between 77 and 82.

3. (5 points) For the distribution of test scores in problem 1, assuming that test scores are only whole numbers, appropriately round your answer up or down to find: **SHOW YOUR WORK**

_____the highest C, if the bottom 43% of the students are given C's, D's & F's.

4. (10 points) On January 28, 1986, the space shuttle Challenger failed shortly after take off. This was the 25th mission in the space shuttle program. Three years earlier, a U.S. Air Force study found that the likelihood of such a failure on any one mission was $\frac{1}{35}$. Assuming that failures are independent from mission to mission, based on the US Air Force's estimate, find the likelihood of:

(Do not round off 1/35th – use the fraction.) **SHOW YOUR WORK**

_____ exactly one failure in 25 missions. (Round your final answer to the nearest ten-thousandth.)

_____ at least one failure in 25 missions. (Round your final answer to the nearest hundredth.)

5. (10 points) For the probability distribution given, find the:

_____ Mean.

_____Standard Deviation.

SHOW YOUR WORK

X	P(X)				
0	0.1				
1	0.3				
3	0.4				
4	0.2				

6. (15 points) Suppose you had a box of 15 identically looking chocolates with 5 being creams and 10 marshmallow. If you select and eat 4 of them at random, find the probability that: **SHOW YOUR WORK**

_____ 3 of them were creams. (Round to the nearest ten-thousandth.)

_____ At most 3 of them are creams. (Round to the nearest ten-thousandth.)

_____ What is the expected number of creams? (Round to the nearest ten-thousandth.)

7. (15 points) Taxi cabs arrive at an average rate of 20 times an hour. If their arrival follows a Poisson distribution, find the likelihood that the number arriving in the **quarter-hour period** between 9:00 AM and 9:50 PM is:

SHOW YOUR WORK

_____ 4.

_____ at least 4.

_____ What is the standard deviation for the bus arrivals?

8. (10 points) Carrie always is equally likely to arrive 5 to 25 minutes late for class each day. If the class starts at 5PM: (Hint: what probability distribution is this problem?)

SHOW YOUR WORK

_____ Find the likelihood that she arrives before 5:20PM.

_____ On average, at what time would you expect her to arrive?

9. (15 points) College students are classified by major and year in school as shown in the table below. If a student is selected at random, find the probability that the student is:

SHOW YOUR WORK

	Art	Math	Science	History	English
Freshman	**10**	**15**	**20**	**10**	**5**
Sophomore	**12**	**6**	**8**	**10**	**6**
Junior	**13**	**15**	**7**	**7**	**6**
Senior	**15**	**14**	**10**	**10**	**1**

_____a Freshman Math student.

_____a Freshman or a Math student.

_____a Freshman, if he or she is a Math student.

Name:_____ **Due:**_____ **Points Earned:**_____

Directions: Answer all questions on these sheets. Written answers must be **printed & legible**. You must use correct grammar and mechanics. Illegible handwriting and/or poor grammar/mechanics will result in the loss of credit.

I. Basic Probability Problems from Chapter 6.

1. (20 points) The passengers on a flight arriving in Baltimore from London are classified as follows:

	U.S. Citizen	European Citizen	Other Citizenship
First Class	20	8	12
Business Class	25	10	15
Coach Class	110	44	66
Tourist Class	75	30	45
Stand-By	20	8	12

If a passenger is selected at random, find the likelihood that the passenger is:

_____A U.S. Citizen.

_____A Tourist Class Passenger.

_____A U.S. Citizen in Tourist Class.

_____A U. S. Citizen or a Tourist Class Passenger.

_____A U.S. Citizen, if we know that the Passenger is in Tourist Class.

_____A Tourist Class Passenger, if he or she is a U.S. Citizen.

_____Not flying Stand-By.

_____Are the classifications U.S. Citizen and Tourist Class Passenger mutually exclusive?

_____Are the classifications U.S. Citizen and Tourist Class Passenger independent?

_____Which probability interpretation did you use here, relative frequency (RF), subjective (SUB), or classical (CLASS)?

2. (12 points) Find the following, if $P(A) = 0.5$, $P(B) = 0.25$ and $P(B|A) = 0.4$.

_____ $P(A \quad B)$.

_____ $P(A|B)$.

_____ $P(A \text{ é } B)$.

_____What must we change the value of $P(B)$ to, in order to make events A and B independent events?

3. (10 points) A pair of fair four-sided dice, numbered 1 to 4, is tossed.

List all possible outcomes._____

_____

_____Find the probability that the sum of the numbers showing is 3.

_____Find the probability that the sum of the numbers showing is not 3.

_____Find the probability that the sum of the numbers showing is odd.

4. (7 points) Little Johnny tossed a bent bottle cap 100 times and on 60 occasions it fell with label facing up. This led him to conclude that the probability of this bottle cap falling with the label facing up is precisely 60.0%. Identify the method Johnny used in determining this probability, and explain whether or not you agree with his conclusion. **Print your answer.**

5. (10 points) Jim and Yvette plan to have a family of two children with girls being less likely than boys. Assume that 52% of all babies born are boys.

List all possible outcomes for this experiment.

_____ Find the probability that all of the children are boys.

_____ Find the probability that at least one of children is a boy.

_____ Which probability interpretation did was used in determining the likelihood of a boy, relative frequency (RF), subjective (SUB), or classical (CLASS)?

6. (9 points) Jack and Jill are inconsistent when it comes fetching a pail of water. On any given day, there is a 60% chance that Jack will fetch a pail of water, and a 40% chance that Jill has fetched one. If neither Jack nor Jill knows whether or not the other one will fetch a pail of water today, find the probability that:

_____ Neither Jack nor Jill fetches a pail of water.

_____ Both Jack and Jill fetch a pail of water.

_____ Only one of them fetches a pail of water.

7. (10 points) Bart Simpson knows that his best strategy, for passing a multiple-choice math test, is to choose the answers at random. After all, he never studies and he knows nothing about the material on the test. If each question had 4 choices (a, b, c, d), find the probability that Bart answers:

_____ The first question correctly.

_____ The first two questions correctly.

_____ The first four questions correctly.

_____ Only the first question correctly if there are four questions.

_____ Does not answer the first four questions correctly.

8. (22 points) The following article appeared in the *New Your Times* on 9/17/96. According to the article, millions of Americans ignore their mother's advice concerning washing their hands after going to the bathroom. Read the article and answer the following questions.

Shame on You! Wash Your Hands! Especially You 40% at Penn Station

NEW ORLEANS, Sept. 16 (AP) — Millions of Americans routinely ignore one of their mothers' most important pieces of advice: wash your hands after you go to the bathroom.

This unsettling item of news was gathered in the only way possible — by actually watching what people do (or don't do) in public restrooms.

The researchers — if that is what they should be called — hid in stalls or pretended to comb their hair while observing 6,333 men and women do their business in five cities last month.

"Hand washing in this country has become all but a lost art," said Dr. Michael T. Osterholm, the Minnesota state epidemiologist.

Dr. Osterholm heads the public health committee of the American Society for Microbiology, which sponsored the survey with the Bayer Corporation, the aspirin maker. The figures were gathered by Wirthlin Worldwide, a survey firm, and made public today at the society's annual infectious-disease conference.

Among the results were these:

¶The country's dirtiest hands may be in New York City. Just 60 percent of those using restrooms in Penn Station washed up afterward.

¶Chicago hands, relatively speaking, are reasonably clean. The watchers saw 78 percent take the time to wash after using the bathroom at the Navy Pier.

¶Seventy-one percent washed up at a casino in New Orleans, 69 percent at Golden Gate Park in San Francisco and 64 percent at a Braves game in Atlanta.

¶Women are cleaner than men. The survey found 74 percent wash after using the toilet, compared with 61 percent of men. Only in New York and New Orleans did men use soap and water slightly more often than women.

¶The most slovenly men observed were at the Braves game. Just 46 percent of the men stopped to wash, compared with 89 percent of women.

Dirty hands are an extremely common means of spreading diseases, ranging from colds to illnesses that cause diarrhea and other intestinal problems. In restaurants, one food handler with dirty hands can make dozens of patrons sick.

People probably know better. Wirthlin also did a telephone survey last month of what people had to say about their hand-washing habits. Of 1,004 adults, 94 percent said they always washed up after using public restrooms.

Actually, mothers' advice — or lack of it — may be part of the problem. "Moms often today are not telling their kids to wash their hands," said Dr. Gail Cassell of the University of Alabama. "Schools aren't telling children about it. We need to be reminded that this is important."

_____ Based on this article the mayor of New York claims that 60% of all New Yorkers was their hands after going to the bathroom. Which probability interpretation did the mayor use, relative frequency (RF), subjective (SUB), or classical (CLASS)?

Explain whether the 60% claim made by the mayor of New York is a parameter or a statistic. **Print your answer.**

The mayor of Old York claims that he is 110% positive that the residents of Old York do a better job of washing their hands after going to the bathroom. Comment on the mayor of Old York's claim from the point of view of a statistician.
Print your answer.

_____ Which probability interpretation did the mayor of Old York use, relative frequency (RF), subjective (SUB), or classical (CLASS)?

When the researchers observed peoples' behavior in the bathroom, they found that they washed their hands after going to the bathroom approximately sixty to seventy-five percent of the time. In a follow up telephone survey, they found that, ninety-four percent responded that they always washed up after using a public restroom. Give a plausible explanation for the discrepancy in the percentages reported by these two surveys.
Print your answer.

Name:_____ **Due:**_____ **Points Earned:**_____

II. Problems from Chapters 6, 7 & 8. (Note: Do not use the Normal Approximation on these problems, since it was not covered in class.)

1. (12 points) With increased airline traffic, it is predicted that in the next century airline accidents will average one every 10 days. If they follow a Poisson distribution, what is the probability that the number of airline accidents in a month (30 days) is:

_____ **a.** at most 7?

_____ **b.** at least 7?

_____ **c.** What is the expected number of airline accidents in 30 days?

_____ **d.** What is the standard deviation of airline accidents in 30 days?

2. (15 points) If it is known that 80% of all Labrador Retrievers are black, in a random sample of 16 Labrador Retrievers, find the probability that the number that are black is:

_____ **a.** 12

_____ **b.** more than 12

_____ **c.** between 12 and 15 inclusive

_____ **d.** What is the expected number of black Labrador retrievers?

_____ **e.** What is the variance for black Labrador retrievers?

3. (9 points) If the life span of Labrador retrievers is known to be normally distributed with a mean of 10.5 years and a standard deviation of 2.1 years. If you own a Labrador retriever, what is the likelihood that it lives:

_____ **a.** at least 14.7 years?

_____ **b.** less than 5.25 years?

_____ **c.** between 5.25 and 14.7 years?

4. (15 points) Assuming independence, find the probability of getting: **Round your answer off to <u>five decimal places</u>.**

_____ **a.** No rain for the next 30 days, if the probability of rain on any particular day is 0.025.

_____ **b.** 8 fours in 10 tosses of a single **<u>four sided die</u>**. (Assume each side is equally likely and numbered 1 to 4.)

_____ **c.** One accident this year (365 days), if the likelihood of an accident on any given day is .005.

_____ **d.** Two or more accidents in a given year, for part "c" above.

_____ **e.** A family of 3 boys and 2 girls, if the likelihood of a boy birth is 0.52.

5. (9 points) A standardized test is known to be normally distributed, with a mean of 890 and a standard deviation of 100. If 96,000 students in a certain region take the test, how many of them would you expect to have a score:

_____ **a.** less than 825?

_____ **b.** at least 1050?

_____ **e.** between 825 and 1050?

6. (6 points) Suppose a binomial distribution has a standard deviation of 12. If the probability of getting a success is p = .40, find:

_____ **a.** the value of n.

_____ **b.** the mean.

7. (6 points) A set of test scores is normally distributed with a mean of 82.4 and a standard deviation of 5.4. Find the following, appropriately rounding your answers up or down, **assuming that each test score is a whole number**:

_____ **a.** the highest B, if the highest 8% are given A's.

_____ **b.** the lowest C, if the bottom 18% are given D's or F's.

8. (6 points) A police officer buys a box of 13 jelly donuts (a baker's dozen). Nine of the donuts are raspberry and four are strawberry. If the police officer randomly selects and eats 5 of the donuts, what is the probability that he or she eats:

_____ **a.** 2 strawberry jelly donuts?

_____ **b.** at least 2 strawberry jelly donuts?

9. (9 points) For the probability distribution given, find the:

_____ Mean.

_____ Variance.

_____ Standard Deviation.

X	P(X)				
2	0.3				
4	0.2				
6	0.1				
8	0.4				

Name:_____ **Due:**_____ **Points Earned:**_____

III. Sampling Distribution and Probability Distribution Problems from Chapters 6, 7, 8 & 9. (Note: Do not use the Normal Approximation on these problems, since it was not covered in class.)

1. (12 points) Suppose that of all students, who took a standardized math test, their average score was 975 with a standard deviation of 100. If test scores are normally distributed, find the probability that a randomly selected student has a test score of:

_____ **a.** at least 945.

_____ **b.** less than 1000.

_____ **c.** between 945 and 1000 inclusive.

2. (12 points) In the previous problem, if we select a random sample of 36 students who took the standardized math test, find the probability that they have an average test score of:

_____ **a.** at least 945.

_____ **b.** less than 1000.

_____ **c.** between 945 and 1000 inclusive.

3. (12 points) At Dog's Breath High School a total of exactly 81 students took a math test. If this test was normally distributed with a mean of 975and a standard deviation of 100, find the probability that a random sample of 36of these students from Dog's Breath High School had an average test score of:

_____ **a.** at least 945.

_____ **b.** less than 1000.

_____ **c.** between 945 and 1000 inclusive.

4. (10 points) Dr. Melinda Marple teaches an Advanced Placement (AP) Calculus to 36 students at Dog's Breath High School (see the previous problem). Suppose that all of her students took the math test, and that their average test score was 1024. Find the probability that a randomly selected group of 36 of the students from Dog's Breath High School had an average math score of at least 1024, and explain whether or not it is reasonable for Dr. Marple's AP Calculus students to score that high on the test.

_____ Probability of an average score of at least 1024. **PRINT YOUR ANSWER**

Explanation:_____

5. (6 points) Dr. Darrel Socrates also teaches 36 math students at Dog's Breath High School. Unlike Dr. Marple's students (see the previous problem), Dr. Socrates' students were of ordinary math ability. Suppose that Dr. Socrates was the person in charge of administering the standardized tests, and that he was the one who received the tests in advance from the company that makes them. Explain what you would conclude, if Dr. Socrates' students had an average score of 1024 on the test. **PRINT YOUR ANSWER**

Explanation:

6. (12 points) Suppose that 12 out of every 100 Labrador Retrievers are chocolate. If 20 Labrador Retrievers are selected at random, find the probability that the number of them that are chocolate is:

a. 5.

b. at least 5.

c. What is the expected number of chocolate Labrador Retrievers in this problem?

7. (8 points) Suppose you find 12 cans of Campbell's Soup scattered all over your cupboard floor with no labels. If 8 of the cans are chicken noodle soup and 4 are alphabet soup, if you select 5 these cans at random and open them, find the probability that the number of cans of alphabet soup that you open is:

a. 2.

b. at most 2.

8. (8 points) Suppose that typos occur at random according to a Poisson Distribution with an average of 1 typo for every 25 pages. Find the probability that the number of typos in an 80-page manuscript is:

_____**a.** 5.

_____**b.** at least 5.

9. (8 points) For the probability distribution given, find the:

_____ Mean.

_____Standard Deviation.

X	P(X)				
1	0.2				
2	0.1				
3	0.4				
4	0.3				

10. (8 points) Sam is always equally likely to be anywhere from 5 to 30 minutes late for class each day. If class begins at 3PM and attendance is completed at precisely 3:12:15PM,

_____**a.** Find the likelihood that he arrives after attendance is completed.

_____**b.** What is his average arrival time?

11. (12 points) Manny, Moe and Jack inconsistent when it comes changing the oil in their cars on time. Manny remembers to change his oil 60% of the time, Moe remembers 40% of the time and Jack remembers a mere 15% of the time. If there is no communication between these three individuals, find the probability that:

_____**a.** None of them remembers to change their oil.

_____**b.** All of them remember to change their oil.

_____**a.** Jack is the only one of them who remembers to change his oil.

MAT 127 – Climent **Test 2 – Take Home Answers** **Pages 1 - 4**

I. 1. (20 points) The passengers on a flight arriving in Baltimore from London are classified as follows:

	U.S. Citizen	European Citizen	Other Citizenship	
First Class	20	8	12	40
Business Class	25	10	15	50
Coach Class	110	44	66	220
Tourist Class	75	30	45	150
Stand-By	20	8	12	40
	250	100	150	500

If a passenger is selected at random, find the likelihood that the passenger is:

250/500 = 0.50 A U.S. Citizen.
150/500 = 0.30 A Tourist Class Passenger.
75/500 = 0.15 A U.S. Citizen in Tourist Class.
325/300 = 0.65 A U. S. Citizen or a Tourist Class Passenger.
75/150 = 0.50 A U.S. Citizen, if we know that the Passenger is in Tourist Class.
75/250 =0.30 A Tourist Class Passenger, if he or she is a U.S. Citizen.
460/500 = 0.92 Not flying Stand-By.
No Are the classifications U.S. Citizen and Tourist Class Passenger mutually exclusive?
Yes Are the classifications U.S. Citizen and Tourist Class Passenger independent?
Class Which probability interpretation did you use here, relative frequency (RF), subjective (SUB), or classical (CLASS)?

2. (12 points) Find the following, if *P(A) = 0.5, P(B) = 0.25* and *P(B|A) = 0.4.*

0.20

0.80

0.55

0.40 What must we change the value of *P(B)* to, in order to make events *A* and *B* independent events?

3. (10 points) A pair of fair four-sided dice, numbered 1 to 4, is tossed.

List all possible outcomes. {(1,1), (1,2), (1,3), (1,4), (2,1), (2,2), (2,3), (2,4), (3,1), (3,2), (3,3), (3,4), (4,1), (2,2), (4,3), (4,4)}

2/16 = 0.125 Find the probability that the sum of the numbers showing is 3.

14/16 = 0.875 Find the probability that the sum of the numbers showing is not 3.

8/16 = 0.500 Find the probability that the sum of the numbers showing is odd.

4. (7 points) Little Johnny tossed a bent bottle cap 100 times and on 60 occasions it fell with label facing up. This led him to conclude that the probability of this bottle cap falling with the label facing up is precisely 60.0%. Identify the method Johnny used in determining this probability, and explain whether or not you agree with his conclusion.

Relative Frequency was the method used in determining the probability of a tossed bottle cap falling with the label facing up. The key point in answering this question was that the relative frequency requires thousands of tosses before one can reliably estimate a probability with reasonable accuracy. Using only 100 tosses cannot reliably give us a precise answer.

5. (10 points) Jim and Yvette plan to have a family of two children with girls being less likely than boys. Assume that 52% of all babies born are boys.

List all possible outcomes for this experiment. **{GG, GB, BG, BB}**

(.52)(.52) = 0.12704 Find the probability that all of the children are boys.

(.52)(.52) + (.48)(.52) + (.52)(.48) = .7696 Find the probability that at least one of children is a boy.

RF Which probability interpretation was used in determining the likelihood of a boy, relative frequency (RF), subjective (SUB), or classical (CLASS)?

6. (9 points)) Jack and Jill are inconsistent when it comes fetching a pail of water. On any given day, there is a 60% chance that Jack will fetch a pail of water, and a 40% chance that Jill has fetched one. If neither Jack nor Jill knows whether or not the other one will fetch a pail of water today, find the probability that:

F = Fetches N = Does not Fetch
P(FF) = (0.60)(0.40) = 0.24
P(FN) = (0.60)(0.60) = 0.36
P(NF) = (0.40)(0.40) = 0.16
P(NN) = (0.40)(0.60) = 0.24

0.24 Neither Jack nor Jill fetches a pail of water.

0.24 Both Jack and Jill fetch a pail of water.

0.52 = 0.16 + 0.36 Only one of them fetches a pail of water.

7. (10 points) Bart Simpson knows that his best strategy, for passing a multiple-choice math test, is to choose the answers at random. After all, he never studies and he knows nothing about the material on the test. If each question had 5 choices (a, b, c, d), find the probability that Bart answers:

1/4 = 0.25 The first question correctly.

(1/4) (1/4) = 1/16 = 0.0625 The first two questions correctly.

(1/4) (1/4) (1/4) (1/4) = 1/256 = 0.00390625 The first four questions correctly.

(1/4) (3/4) (3/4) (3/4) = 27/256 = 0.10546875 Only the first question correctly if there are four questions.

1 – 1/256 = 255/256 = 0.99609375 Does not answer the first four questions correctly.

8. (22 points) According to the accompanying article, millions of Americans ignore their mother's advice concerning washing their hands after going to the bathroom.

RELATIVE FREQUENCY Based on this article the mayor of New York claims that 60% of all New Yorkers was their hands after going to the bathroom. Which probability interpretation did the mayor use, relative frequency (RF), subjective (SUB), or classical (CLASS)?

Explain whether the 60% claim made by the mayor of New York is a parameter or a statistic. **It is a statistic, because it is based on a sample of New Yorkers. To be a parameter it would have to be based on all New Yorkers.**

The mayor of Old York claims that he is 110% positive that the residents of Old York do a better job of washing their hands after going to the bathroom. Comment on the mayor of Old York's claim from the point of view of a statistician. **The key point here is that 110% cannot be a probability, since by definition all probabilities must be between 0 & 1 (0% & 100%). Another point is that the estimate appears to be subjective and not based on any objective data.**

SUBJECTIVE Which probability interpretation did the mayor of Old York use, relative frequency (RF), subjective (SUB), or classical (CLASS)?

When the researchers observed peoples' behavior in the bathroom, they found that they washed their hands after going to the bathroom approximately sixty to seventy-five percent of the time. In a follow up telephone survey, they found that, ninety-four percent responded that they always washed up after using a public restroom. Give a plausible explanation for the discrepancy in the percentages reported by these two surveys. **People are probably too embarrassed to admit the truth to other people. Perhaps they think the always wash their hands after they go to the bathroom, but in reality they don't. Perhaps the restrooms are too dirty to wash in.**

Directions: 1. Show all work on these sheets.
2. Calculators are permitted.
3. Sit in assigned seat
4. No Sharing of materials with others.
5. You may use the Minitab printouts and calculator work you prepared beforehand for this test.
6. You may use three pages of hand written notes and the handouts given in class.
7. You may use a photocopy of all of the tables in the back of your book.
8. You may not use your textbook and you may not your notebooks.
9. Make sure your symbols are written clearly: **μ, _, s, p,** etc.

Name:_____

Grade = _____ = _____ %.
150

I. (112 points) Problems prepared in advance.

1. (14 points) Problem 1: Is there sufficient evidence to conclude that in-state students score higher on the math SAT test than out-of-state students? If our level of significance is specified as $\alpha = 0.05$, answer the following questions and fill in the information requested.

_____
Smallest p-value for the normality assumption

_____ _____
Value of the correct computational formula. Write out the correct computational formula.

_____H_1 for testing the **SAT scores.**

Reject H_0 if _____Critical Region for testing the **SAT scores.**

_____Is there sufficient evidence to conclude that in-state students score higher than out-of-state students?

_____What is the value of the sample mean for the out-of-state students?

2. (4 points) Problem 2: Identify the major possible bias and describe an experimental design that could avoid this bias.

Identify the bias **(PRINT)**:_____

_____

Describe a better experimental design **(PRINT)**:_____

_____

_____

3. (14 points) Problem 3: Is there sufficient evidence to conclude that nontraditional students perform better in statistics than traditional students? If our level of significance is specified as $\alpha = 0.05$ fill in the requested information.

_____Smallest of the p-values for testing **Normality.**

_____H_1 for the **equal variances assumption.**

_____Value of computational formula for testing the **equal variances assumption.**

Reject H_0 if _____Critical Region for testing the **equal variances assumption.**

_____H_1 for testing the **means or location.**

_____Value of computational formula for testing the **means or location.**

Reject H_0 if _____Critical Region for testing the **means or location.**

4. (12 points) Problem 4: Does pay based on commission make the sales force more productive than pay based on hourly wages? Use $\alpha = 0.01$ to answer the following.

_____
Smallest p-value for the normality assumption

_____ _____
Value of the correct computational formula. Write out the correct computational formula.

_____H_1 for testing the **pay effect.**

Reject H_0 if _____Critical Region for testing the **pay effect.**

Conclusion stated in terms of the Pay Levels: **(PRINT)** _____

_____

_____

5. (18 points) Problem 5: Is there sufficient evidence to support the dean's belief? Using $\alpha = 0.10$, answer the following questions and fill in the information requested.

_____H_0

$E(n_1) =$_____ $E(n_2) =$_____ $E(n_3) =$_____ $E(n_4) =$_____

_____Value of computational formula.

Reject H_0 if _____Critical Region.

_____The p-value of the statistic is closest to which of the following numbers: 0.10, 0.05, 0.025, 0.01, 0.005.

_____ < _____ < _____Confidence Interval for seniors (place values and the symbols in the inequalities.)

6. (6 points) Problem 6: Is there sufficient evidence to conclude that the planned commutation modes are proportionally different across the four job categories? If our level of significance is specified as $\alpha = 0.01$, answer the following questions and fill in the information requested.

_____H_0

_____Value of computational formula.

Reject H_0 if _____Critical Region.

II. (38 points) Problems given on test (not given beforehand).

10. (6 points) Suppose that we are testing H_0: $\mu = 75$ and it is appropriate to use the **z-test**. If the numerical value of our statistical formula is $z = -1.78$, find the **p-value** that corresponds to the following possible alternative hypotheses.

_____H_1: $\mu < 75$

_____H_1: $\mu > 75$

_____H_1: $\mu \neq 75$

11. (4 points) For the Anderson-Darling Normality test, state the null and alternative hypotheses.

_____ _____
null hypothesis alternative hypothesis

13. (2 points) What assumption must be satisfied when using the χ^2 distribution to test for proportions or independence?

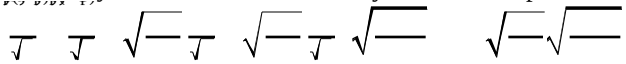
Assumption: _____

MAT 127 **Final – Fall 2002 - Take Home Problems** **John Climent**

Approximately 3/4 of your Final will consist of questions concerning the problems given on the following sheets. The remainder of your test may involve several general questions that require some minimal computations and looking up values in statistical tables.

Instructions:

1. Be prepared to answer questions concerning the following problems and their accompanying data sets.
2. You may use Minitab or a calculator to answer these questions.
3. The data sets are available in Minitab form from your instructor in class. If you fail to get them from him, you will have to enter the data onto Minitab by hand from the printed copies on the other side of these sheets.
4. For each problem, make sure you prepare in advance the appropriate summary statistics.
5. For each problem you will be asked to choose the most appropriate statistical test or procedure to use from those discussed in class.
6. For each problem, you may be asked to describe a suitable population based on the data given and the narrative of the problem. In addition, you should be prepared to address if the study is observational or a designed experiment.
7. For each problem, you should be prepared to explain whether or not the sample is a random sample. If it is not a random sample, you should be prepared to explain why it is or is not representative of the population. If it is not representative, prepare examples of possible biases.
8. For hypothesis testing problems you should be prepared to supply the usual seven pieces of information: H_0, H_1, level of significance, correct statistical formula, value of the statistical formula, the decision rule in terms of both the **critical region and p-values**, and your conclusion written in a narrative in the context of the problem. In addition, for two-tailed tests you should prepare confidence intervals where appropriate.
9.
10. If the level of significance is not specified on any of the problems listed on these sheets, it will be given to you on the actual test. This means you may have to look up the critical region on the test. It also means that your conclusion as to whether or not to reject the null hypothesis may depend on the significance level specified on the test.
11. For confidence interval problems you should be prepared to state the formulas, values of the formulas, the critical value of the appropriate statistic and state the confidence interval in both inequality and interval formats. In addition, you should be prepared to find the size of the sample needed for greater precision.
12. In all of your statistical formulas, you will be expected to give the **complete formula for the standard deviation, e.g.,**



13. Be prepared to illustrate your results with appropriate graphs, including where appropriate, box plots (single and simultaneous), stem-and-leaf diagrams (individual and back-to-back) and histograms. Be prepared to describe the overall shape of these graphs, identify any outliers and describe any important features.
14. If you use a **totally inappropriate statistical test,** then you will l**ose all credit** for that particular problem. Otherwise partial credit is available. For example, if you use a two-sample t-test, when the paired-observations t-test is called for, then you will lose all credit for that problem. If your error is less severe, such as using the z-test, when the pooled-variance t-test is the correct test, then partial credit may be granted for correct answers provided they match the statistical test you used.
15. You will be **permitted to use** the following materials on the exam
    a. The entire Minitab or calculator outputs you prepared in advance.
    b. Two page of handwritten notes.
    c. Your instructor's handout on one and two sample tests.
    d. Copies of the tables in the back of the book. **Make sure you make copies of them and bring them with you.**
    e. Calculators.
12. You **cannot use** the following materials on the exam:
    a. You may not use your textbook.
    b. You may not use your notebooks.
    c. You may not use copies (handwritten or Xeroxed) of previous tests from the current semester or previous semesters.
    d. Shared materials or calculators with other students.

MAT 127 **Final – Fall 2002 - Take Home Problems** **John Climent**

Problem 1:
A dean at Bald Eagle State University, a state run public university, wanted to compare the SAT math scores for in-state and out-of-state students. The dean compiled data from a large elementary education class with 82 students recording the math SAT scores for each student in the class. The data is shown below and it can be found on the Minitab file: **'O2Fall-MAT127-Final-SAT Scores'**. If the level of significance is specified as 0.05, is there sufficient evidence to conclude that in-state students score higher on the math SAT test than out-of-state students?

Math SAT Scores for Out-of-State Students							**Math SAT Scores for In-State Students**						
580	468	597	636	551	577	584	419	621	560	382	348	534	466
548	564	547	527	553	476	483	609	506	455	570	436	543	556
562	577	494	578	545	574	476	485	500	467	666	565	492	550
650	583	610	526	619	638	560	386	376	350	526	628	538	528
688	526	718	554	587	590	594	598	590	549	578	557	628	549
719	592	589	497	505	643		553	430	604	502	617	552	

Problem 2:
There is at least one major possible bias in the design of the experiment in the previous problem. Be prepared to explain in writing what it is, and describe an experimental design that could avoid this bias. Keep in mind that your description of the target population should reflect the dean's intentions.

Problem 3:
Do nontraditional students perform better in statistics than traditional ones? A group of nontraditional and traditional students took the same stat class and their semester averages were recorded. At a level of significance of 0.05, is there sufficient evidence to conclude that nontraditional students perform better than traditional students. This data can be found on the Minitab Worksheet titled **'02Fall-MAT127-Final-Stat Grades'**. Note: the data has been sorted for your convenience.

Students Performance in Statistics													
Nontraditional	71	76	78	80	81	82	85	86	88	92	94	97	99
Traditional	58	61	63	66	67	71	74	78	83	87	92	95	102

Problem 4:
Does pay based on commission make the sales force more productive than pay based on hourly wages? To see if this is true a manager compared each worker's productivity under both compensation plans. The data is shown in the table below, and it can be found on the Minitab Worksheet titled: **'02Fall-MAT127-Test1-Pay'**. At _ = 0.01, is there sufficient evidence to conclude that commission based pay improves productivity?

Productivity of Workers																	
Worker	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**
Hourly	80.5	74.9	76.4	78.0	73.9	79.1	77.6	67.6	82.7	74.1	66.2	83.9	76.5	74.1	73.0	78.1	70.9
Comm.	81.4	79.5	79.7	79.6	73.3	77.2	76.4	71.1	84.6	74.5	69.2	84.0	73.1	79.7	76.2	78.6	76.3

Problems 5: A college dean believes that freshmen and sophomores are three times more likely than juniors or seniors to take a certain English instructor's class. This instructor is a very hard grader and the dean feels that juniors and seniors tend to avoid classes taught by this instructor. This semester there are 228 students enrolled in this instructor's classes and the number of freshmen, sophomores, juniors and seniors is 96, 74, 40 & 18 respectively. At a level of significance **_ = 0.10**, is there sufficient evidence to support the dean's belief. In addition, find the approximate p-value for your statistic and find a 90% confidence interval for the seniors.

Problem 6: A company is relocating to a new location. They poll their employees as to the how they plan to commute to the new location. The results of the poll are shown in the table to the right. At **_ = 0.01**, is there sufficient evidence to conclude that the planned modes of commutation are proportionally different across the four job categories?

	Car	Bus	Train
Maintenance	40	72	44
Support	90	140	60
Management	18	13	7
Executive	12	4	6