





Stats 2B03 Su25 Assignment #1

« Back to Admin Page

Assignment #1 for: AKSHITA KUMAR (400509297)

### **Instructions:**

- 1. Save your work occasionally (at least once every 15 minutes) or your session will be timed out and you will lose any unsaved work.
- 2. You can edit and change your answers as many times as you want before the deadline.
- **3.** You can submit your work for grading a maximum of 3 times. (If you save your work but do not submit it for grading then it will automatically be submitted after the deadline.)

Topics: Graphical Displays and Descriptive Statistics Sections Covered: 1.1, 1.2, 2.1-2.3, 3.1-3.3

igwedge denotes questions that require the use of R. (R should NOT be used for any of the other questions.)

If a question requires R, then the answer for that question should be copied and pasted directly from R into the answer box (whenever possible), without doing any rounding.

If a question requires the use of R and the question also specifies that the answer should be given to a specific number of decimals, then you can still copy and paste your answer directly from R (without rounding), if you want. Or you could round the answer to the specified number of decimals. Either way is okay in that case.

**Problem #1:** Read the above instructions and then answer the following questions.

- (a) For questions that require the use of R, how many decimals should be used for the answer?
- **(b)** What if a question requires the use of R, but also specifies that the answer should be given correct to 3 decimals (for example)?
- (A) Exactly 4 decimals.
- **(B)** The output should be copied and pasted directly from R into the answer box, without doing any rounding.
- (C) Exactly 3 decimals. (D) At least 3 decimals. (E) At least 4 decimals.

Problem #1(a): B ➤

(A) The output can be copied and pasted directly from R into the answer box, without doing any rounding. Or the answer could be rounded to the given number of decimals. Either way is okay in that case.

- **(B)** The answer MUST be given to exactly 3 decimals.
- (C) The answer MUST be copied and pasted directly from R

Problem #1(b): A ✓

Just Save Submit Problem #1 for Grading

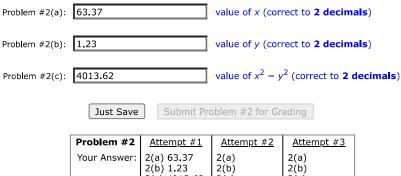
Problem #1	Attempt #1	Attempt #2	Attempt #3
Your Answer:	1(a) B	1(a)	1(a)
	1(b) A	1(b)	1(b)
Your Mark:	1(a) 1/1 <b>√</b>	1(a)	1(a)
	1(b) 1/1 <b>√</b>	1(b)	1(b)

## **Problem #2:** This problem is a note about rounding that you should keep in mind when doing any calculations in this course.

- (a) In most circumstances, numbers are rounded in the usual way. (Exceptions to this will arise in a few places during the course.) So, for example, if you are asked for an answer correct to 2 decimals, then an answer like 3.245873 would be rounded to 3.25.
  - Suppose that you calculate the value of x to be 63.3651724. Enter the value of x into the answer box, correct to 2 decimals.
- (b) Suppose that you calculate the value of y to be 1.23481367. Enter the value of y into the answer box, correct to 2 decimals.
- (c) Now suppose that you need to use the preceding x and y values for another calculation. Then you must use the unrounded versions of those numbers. If you use the rounded the versions of x and y then your answer will be subject to rounding error, and will most likely be wrong.

With this in mind, find the value of  $x^2 - y^2$ , and enter it into the answer box.

Then save your work and submit it for grading.



Problem #2	Attempt #1	Attempt #2	Attempt #3
Your Answer:	2(a) 63.37 2(b) 1.23 2(c) 4013.62	2(a) 2(b) 2(c)	2(a) 2(b) 2(c)
Your Mark:	2(a) 1/1 <b>√</b> 2(b) 1/1 <b>√</b> 2(c) 1/1 <b>√</b>	2(a) 2(b) 2(c)	2(a) 2(b) 2(c)

Problem #3: Using the project kumaa25\_class.RData that you created in Assignment #0, verify using the sum of the command with the na.rm=T option (similar to the mean command used in Assignment #0) that the sum of the height column is equal to 24158.

(Note that this check is to make sure that you are using the correct personalized class data set, and that the data used for this question has not been inadvertently altered.)

Work through this example on R, and then do the following:

Using the project kumaa25 class.RData that you created in Assignment #0, make a grouped frequency distribution similar to this example using the height variable (height) and the intervals 0-61, 61-64, 64-67, 67-70, 70-73, 73-100

Do not submit the grouped frequency distribution, but instead answer the following questions based on the grouped frequency distribution.

(a) Enter the values from the CumPct column of your grouped frequency distribution (in order from top to bottom) separated by commas. For example, if your grouped frequency distribution was the same as at the bottom of this example then you would enter the following numbers into the answer box (note that spaces don't matter):

(b) What percentage of the heights (for which the height data was available) are at least 70 inches?

values from CumPct column. Problem #3(a): 1.923077,24.725275,55.769 separated by commas Enter your answer as a percentage, correct to 2 decimals, Problem #3(b): 22.25 without the % sign. e.g., 28.31 Just Save Submit Problem #3 for Grading Problem #3 Attempt #1 Attempt #2 Attempt #3 3(a) 1.923077, Your Answer: 24.725275 55.769231, 77.747253. 3(b) 15.66 3(b) 22.25 93.406593. 100.000000 3(b) 77.77 Your Mark: 3(a) 6/6 < 3(a) 3(a) 3(b) 0/2X 3(b) 2/2 ✓

3(b) 0/2X

Problem #4: Using the project kumaa25\_class.RData that you created in Assignment #0, verify (using the sum of command with the na.rm=T option) that the sum of the study column is equal to 1396.6, and that the sum of the gender column is equal to 527.

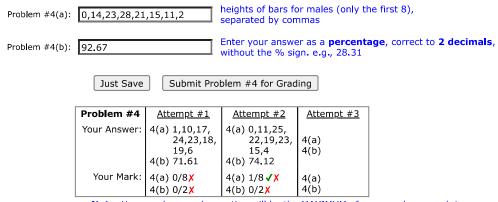
> (Note that this check is to make sure that you are using the correct personalized class data set, and that the data used for this question has not been inadvertently altered.)

Work through this example on R, and then do the following:

Using the project kumaa25 class.RData that you created in Assignment #0, make histograms (2 in total) for the time spent studying (study) for each gender (the gender variable gives the gender, 1 = female, 2 = male) using the intervals 0-1, 1-2, 2-3, 3-4, 4-5, 5-6, 6-7, 7-8, 8-9, 9-10. Make sure that the tickmarks on the xaxis are at the cutpoints and be sure to label the heights of all of the bars.

Do not submit the histograms, but instead answer the following questions based on the histograms.

- (a) Enter the first eight values of the heights of the bars for males (in order from left to right) separated by commas. For example, if your histograms looked like the histogram for males in this example then you would enter the following numbers into the answer box (note that spaces don't matter): 0, 1, 9, 17, 12, 8, 2, 1
- **(b)** What percentage of the females study at least 3 hours per day, on average?



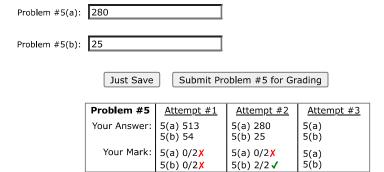
Note: Your mark on each question will be the MAXIMUM of your marks on each try. (So there is no harm in making another attempt at a partially correct answer.)



Using the project kumaa25 class.RData that you created in Assignment #0, convert the time spent using your cell phone in minutes (cell) to hours by dividing by 60 (note that 1 hour = 60 minutes). Verify that the mean of the new hours column is equal to 3.26152.

Construct a stem-and-leaf plot of these times in **hours** using intervals of length 1, and then answer the following questions **based on the depths** of the stem-and-leaf plot.

- (a) How many people spend at least 2 hours per day using their cell phone?
- **(b)** How many people spend at least 1 hour, but less than 2 hours using their cell phone?

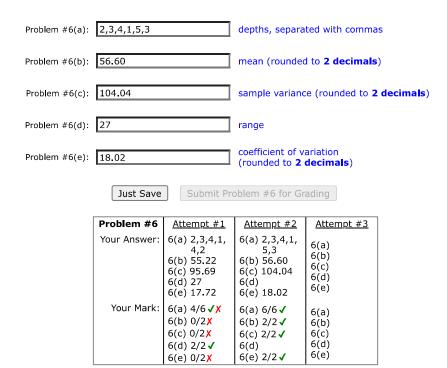


# **Problem #6:** Note that this question (along with all other problems that don't have the R symbol beside them) should be done by hand.

Consider the following data,

(a) Draw a stem and leaf plot **including depths** using intervals of length 5. (See the bottom of <u>this example</u> for an explanation of depths.) Enter the value of the depths **only** in the answer box below. For example, if your stem-and-leaf plot looked like the one at the bottom of <u>this example</u>, then you would enter the following numbers into the answer box:

- **(b)** Calculate the mean.
- (c) Calculate the sample variance.
- (d) Calculate the range.
- (e) Calculate the coefficient of variation.



**Note:** Your mark on each question will be the MAXIMUM of your marks on each try. (So there is no harm in making another attempt at a partially correct answer.)

## **Problem #7:** Consider the data set that is summarized in the R Output below.

- (a) Find the values of  $Q_1$  and  $Q_3$ .
- (b) Find the median
- (c) Find the adjacent values.

(Note: See this example for the relevant definitions and an example.)

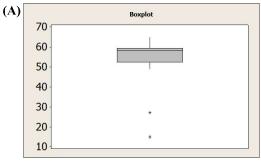
**(D)** 

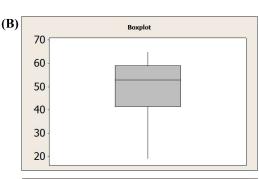
(d) Which of the following is a correct modified boxplot for this data set?

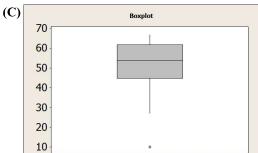
Problem #7(a): 42,58 separate your answers with a comma (the order doesn't matter)

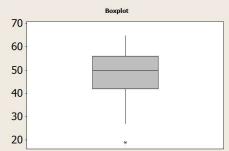
Problem #7(b): 51 median

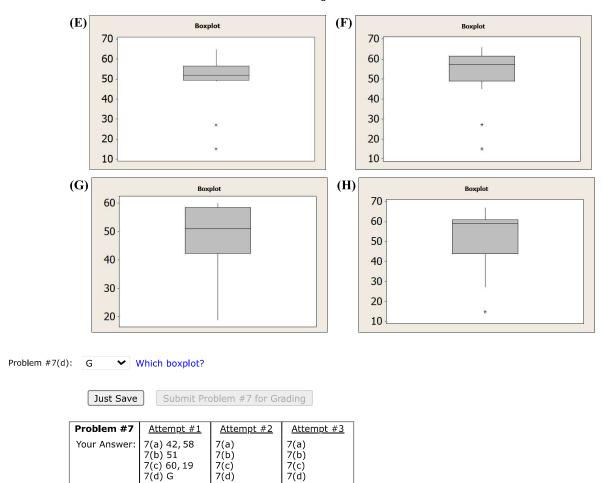
Problem #7(c): 60,19 separate your answers with a comma (the order doesn't matter)











7(a) 7(b) 7(c) 7(d)

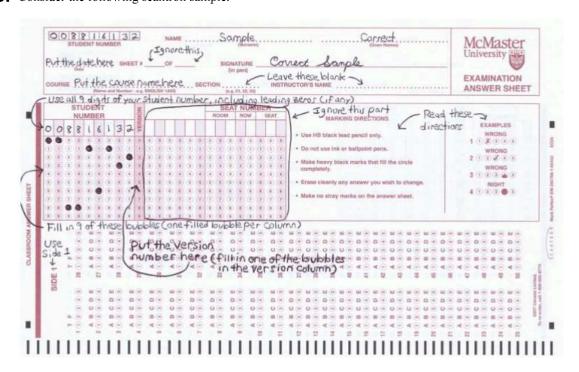
**Problem #8:** Consider the following scantron sample.

Your Mark:

7(d) G

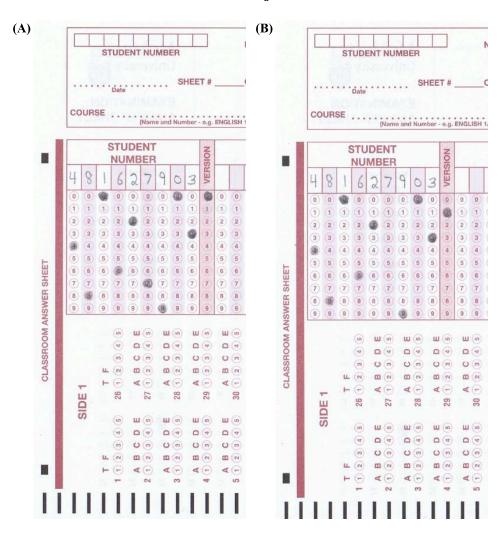
7(a) 2/2 ✓

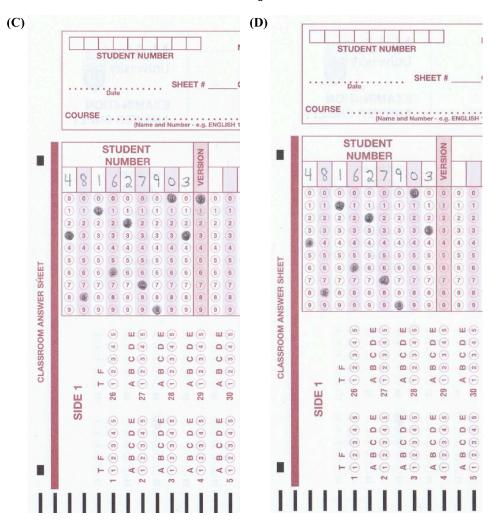
7(b) 2/2 **√** 7(c) 2/2 **√** 7(d) 2/2 🗸 7(a) 7(b) 7(c) 7(d)

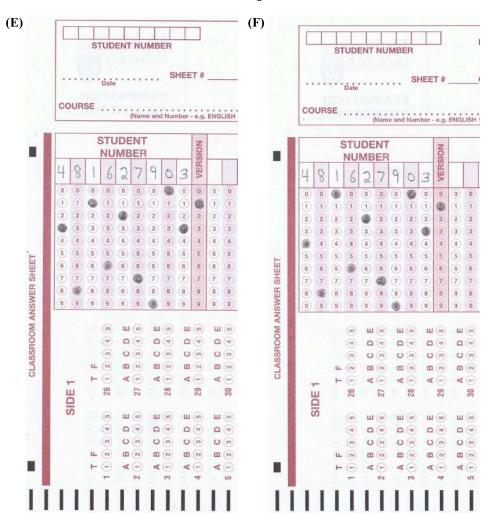


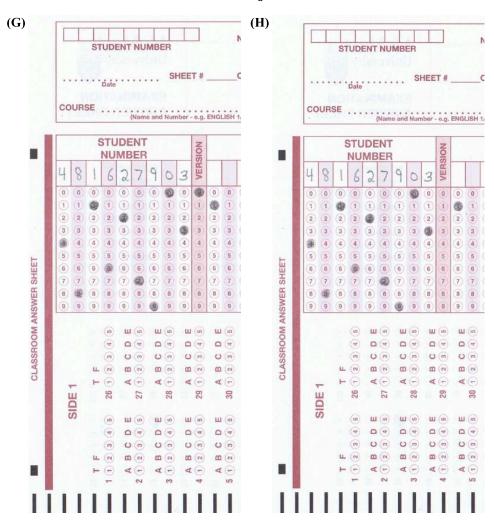
Suppose that a person's student number is 481627903 and that they are writing Version 1 of the test. Which of the

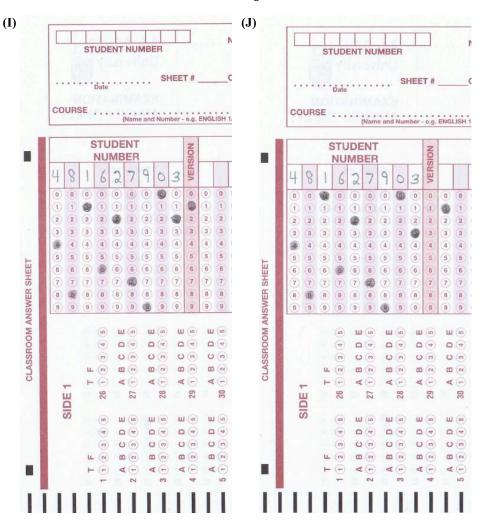
78/25, 10:17 PM	Stats 2803 Su25 Assignment #1 Answer Sheet for Oser: kumaa25
	following scantrons has the bubbles corresponding to the both the student number and version number of the test correctly filled out on the scantron?
	correctly filled out on the scantron?
	contour, miles can on the scanner.
T.	

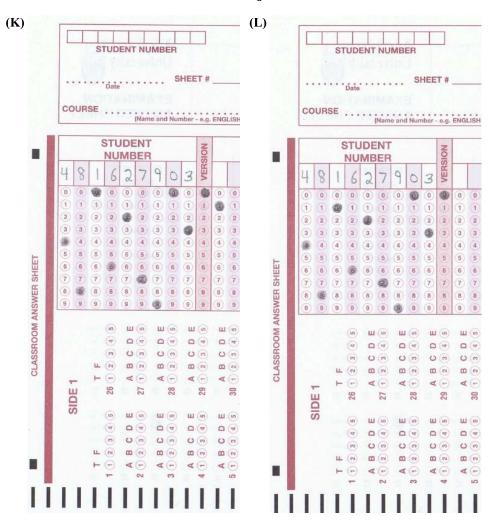


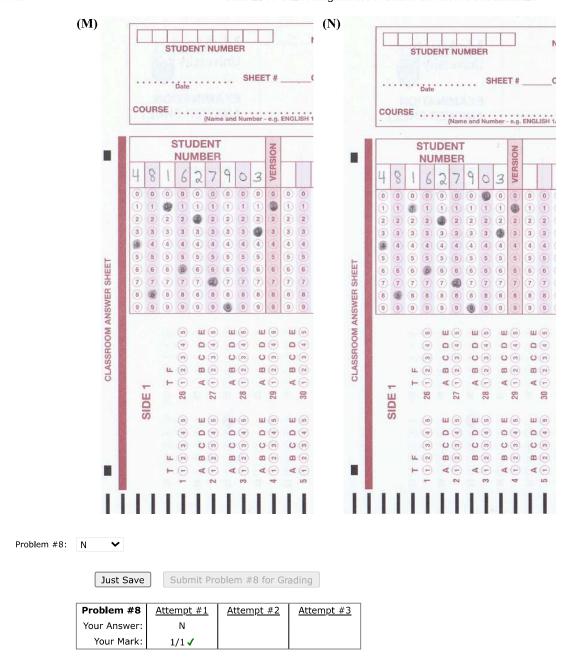




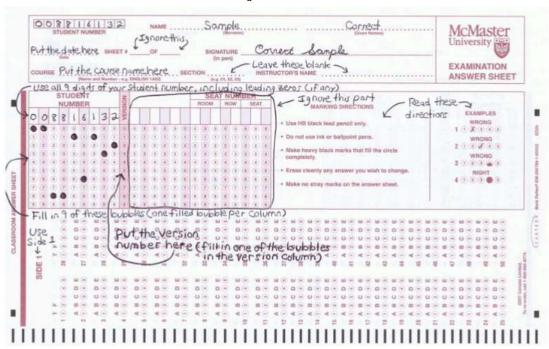




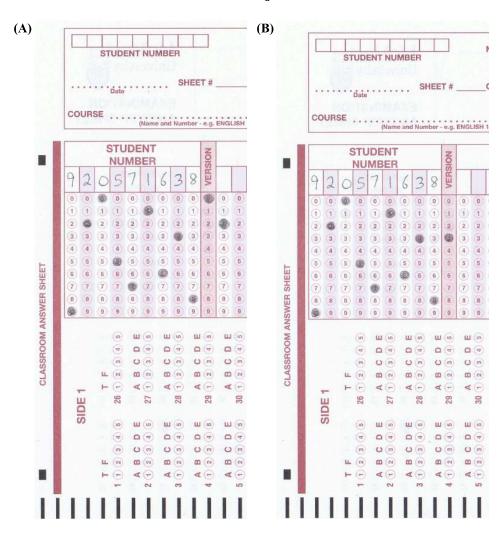


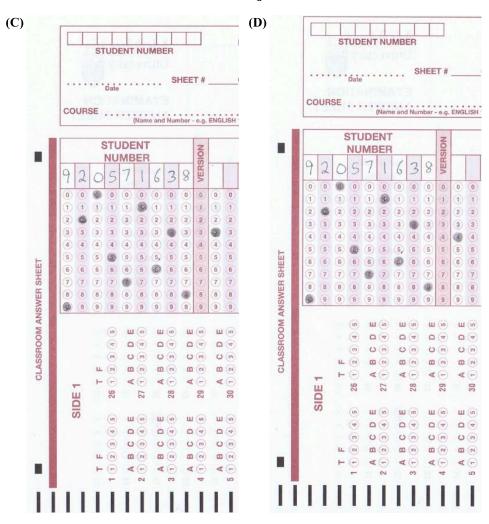


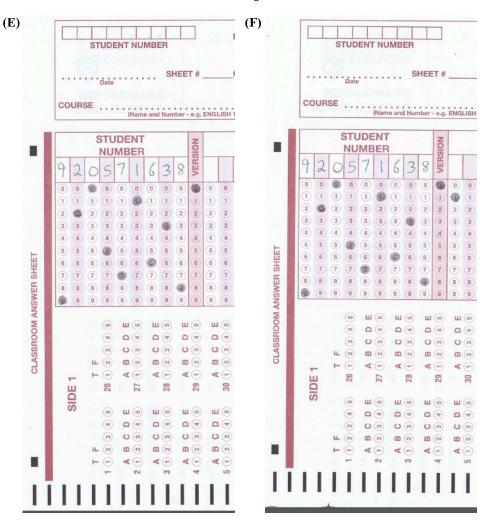
**Problem #9:** Consider the following scantron sample.

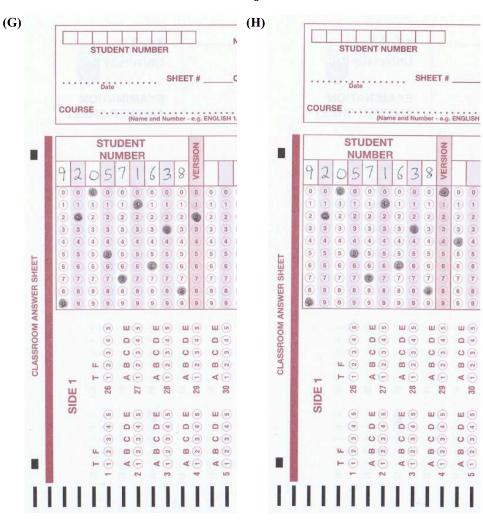


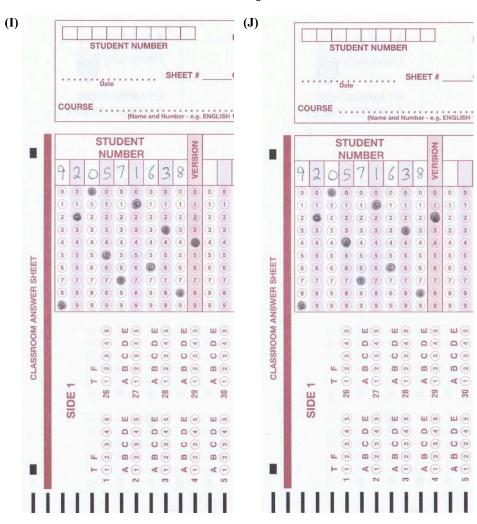
Suppose that a person's student number is 920571638 and that they are writing Version 4 of the test. Which of the following scantrons has the bubbles corresponding to the both the student number and version number of the test correctly filled out on the scantron?

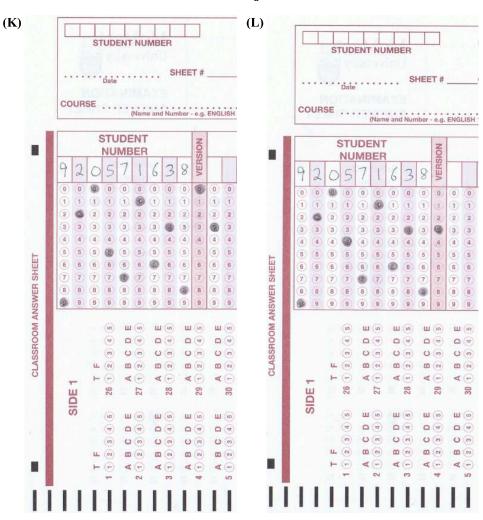


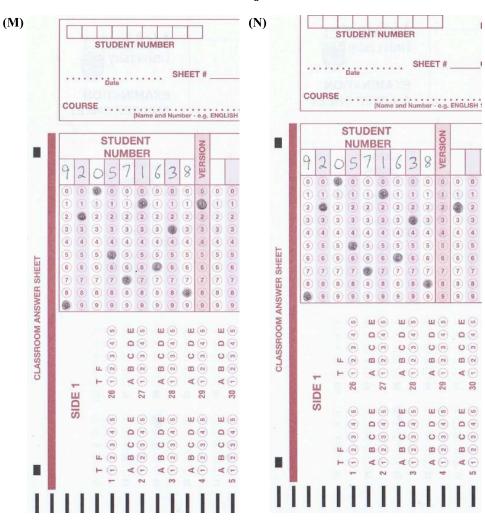


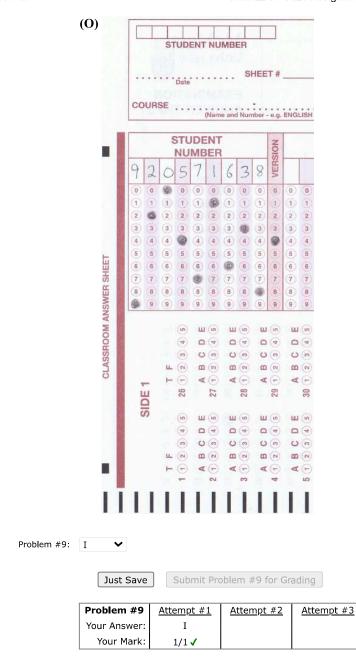












**Problem #10:** On the first test the last question will be worth 1 mark and will require you to correctly fill in the bubbles corresponding to your student number and the version number of your test on the scantron. If you fail to do either of these things correctly then you will not get the mark for this question. Suppose that the test is out of a total of 21 marks (including the one mark for the last scantron question).

How many marks will you lose on your test if you fail to correctly fill in the bubbles corresponding to your student number and the version number of your test on the scantron?

(A) 1.71% (B) 1% (C) 5% (D) 2.36% (E) 2.88% (F) 5.26% (G) 4.76% (H) 0.56%

