

Math 263: Excel Assignment 5 (Version 2)

You must do the computer work for this assignment yourself, although you may certainly talk to other people. Graphs or answers which appear to be copied will be treated as an integrity violation. In particular, everyone needs to create his/her own simulation.

Topic: Test Taking Patterns by Students Applying to College that Require the SAT

Students that apply to colleges that require the ACT and SAT and students that apply to graduate programs that require the GRE, MCAT, or GMAT often take the test more than once, hoping for a better score.

Currently, admissions offices in colleges requiring the SAT only use a student's highest score. This arguably favors wealthier students, who can afford to take the test more than once.¹ Proposals have been made that students who take the SAT more than once should be judged on their

- Maximum score (the current policy), or
- Most recent score, or
- Average score

We will look at simulated data to see the effect of taking two and five tests on these three measures. A student's score may change from one test to the next by random variation—because the student has “a good day” or a “bad day”. Taking the test repeatedly may also increase scores because a student gets practice. Our simulations look at the effect of random variation on the three measures, but not on the effect of practice.

We consider a student whose scores, if he/she took the SAT repeatedly, would be normally distributed with mean 1200 and standard deviation 30.

¹ See the article in *Chronicle of Higher Education*, Jan 29, 2003, summarizing “College Try: Why Universities Should Stop Encouraging Applicants to Take the SATs Over and Over Again” on American Prospect web site. Reported in *Statistics* 7th ed, by R. Peck and J. Devore.

TAKING THE SAT TWICE**Simulation On the PC:**

In the spreadsheet *SAT-Repeats.xlsx*, on the pages *Two Tests*, generate two columns of data with 1000 entries in each. (Put in the pale yellow shaded area.) Do this using the Random Number Generator in the Data Analysis Toolkit, filled out like this:

How does this dialog box work?

- Number of variables is 2 because you want two columns of numbers.
- Number of random numbers is 1000 because we want 1000 numbers in each column.
- Mean is 1200, the student's mean score.
- Standard deviation is 30, the student's standard deviation.
- Output is cell B5 where you want top of the data left of the data to go.

What does the data represent? Each row is a possible pair of values if the student takes the SAT twice.

- In Column D, find the maximum score for that row using $=\text{MAX}(\text{array})$.
- In Column E, find the average score for that row using $=\text{AVERAGE}(\text{array})$.
- In Rows 1 and 2, find the mean and standard deviations of each of the three measures. (Note that the most recent scores are in Column C.)

Simulation On the Mac:

- Open StatPlus
- Click Data at the top of your screen
- Move your cursor down to Random Number Generator and then click Normal Distribution
- The window below will open up.

How does this dialog box work?

- The number of new variables is 2 because each student takes two tests.
- Random Numbers Count is 1000 because we have 1000 students in this experiment.
- Mean is 1200, the mean of the student's scores.
- Standard Deviation is 30, the standard deviation of the student's score.

After hitting OK

- This will pull up a second spreadsheet with 2 columns and 1000 rows of random numbers.
- Copy/Cut and Paste all of these numbers into *SAT-Repeats.xlsx* given with the assignment.
- WARNING! Your data may reappear as ##### in the cell. To correct this, change the number of decimal places of all your cells. This should correct the error.

What does the data represent? Each row is a possible pair of values if the student takes the SAT twice.

- In Column D, find the maximum score for that row using =MAX(array).
- In Column E, find the average score for that row using =AVERAGE(array).
- In Rows 1 and 2, find the mean and standard deviations of each of the three measures. (Note that the most recent scores are in Column C.)

TAKING THE SAT FIVE TIMES**Simulation on Both PC and Mac**

In the same spreadsheet, complete the page *Five Tests* in the same way as you did for two tests. All that needs to change is the Number of (New) Variables.

Note: The page *Five Tests* is in the same Excel file, except it is in a different tab located at the bottom of the spreadsheet.

QUESTIONS TO ANSWER: Comparing the Results of Two and Five Tests

1. Fill in the table below with the results of your simulations.

<i>For 1000 Simulated Values</i>	Maximum Score		Average Score		Most Recent Score	
	<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>
Two Tests						
Five Tests						

Comment on the results:

2. What is effect (if any) of taking the test more often on the mean?
 - (i) For the Maximum score
 - (ii) For the Average score
 - (iii) For the Most Recent score

In each case, explain intuitively why what you observed is to be expected.
3. What is effect (if any) of taking the test more often on the standard deviation?
 - (i) For the Maximum score
 - (ii) For the Average score
 - (iii) For the Most Recent score

In each case, explain intuitively why what you observed is to be expected.
4. Explain how the Central Limit Theorem relates to what you observe for the Average scores.
5. Which measure do you think colleges should use? Why? (There is no one correct answer to the question, but please justify the measure you choose.)