

## Final Project Write-up

### 2.2

Nomen Global is a privately-run English Language Center, and has a purpose similar to that of Brigham Young University's English Language Center: preparing students from non-English language backgrounds for English-based communication in both a general and collegiate-academic sense.

Nomen Global's curriculum centers on a theoretical and proprietary program of development, which differs from the traditional government –approved style of training. As a result of its divergence from this industry standard, Nomen Global has had to fight for accreditation.

The company founder, a friend of mine, approached me in August 2013 requesting that I help with looking at the data Nomen Global had compiled on student performance. I have been working on a part-time basis to help Nomen Global justify its divergent program to the appropriate government agency. At the end of January 2014, I learned Nomen Global received a one-year trial accreditation based on some of the calculations I've helped them perform in VBA; furthermore, this (my final) project is being developed in an effort to help Noman Global acquire more substantial accreditation approval.

In my project, I have worked with three different sets of information for Nomen Global. Two of the data sets contain student performance in specific course work, and one set of data holds student exam performance (each datum compares a student's exam performance at the start of education and following a substantial amount of education).

The two sets of data regarding students' course-work performance, although related, are from two different sources; therefore, I had to treat them quite differently when writing code because the arrangements of the data were significantly different (in fact, I'm still somewhat struggling in dealing with the older of the two datasets, but have made some progress).

For comparing each student's exam performances and the *new* course-work performances, I wrote VBA code that looped through the student names and accessed the appropriate cells relative to the location of the student's name. The reason for the simplicity of this part is the fact that I personally exported the information from their current database in ".csv" format and was able to arrange this data in Excel files in a neat format that works well for iterative processing. The simplicity of the code needed for this is conveyed by how I arranged the data: I emphasized consistent spacing and, because I personally placed the data in these files, I knew where the important information was located on each sheet.

*\*See the following screen captures which are included to emphasize the arrangement of the two datasets (student names have been blacked out). The first Screen capture is for the exam comparisons and the second is for the new course-work performance calculations.*

Last Name	First Name	ID Number	Grammar Test	Listening Test	Reading Test
		5/13/2013	Grammar Test	506.7	Listening Test
		8/13/2013	Grammar Test	47.20	Listening Test
		6/5/2013	Grammar Test	636.1	Listening Test
		8/13/2013	Grammar Test	735.2	Listening Test
		7/15/2013	Grammar Test	99.10	Listening Test
		8/13/2013	Grammar Test	591.1	Listening Test
		6/5/2013	Grammar Test	100	Listening Test
		8/13/2013	Grammar Test	378.3	Listening Test
		6/5/2013	Grammar Test	401.8	Listening Test
		8/13/2013	Grammar Test	560.2	Listening Test
		4/29/2013	Grammar Test	395.4	Listening Test
		8/13/2013	Grammar Test	525	Listening Test

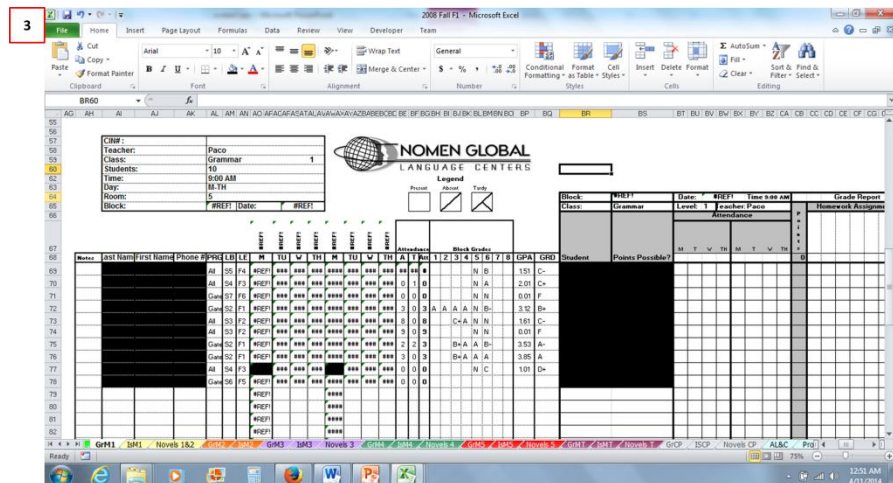
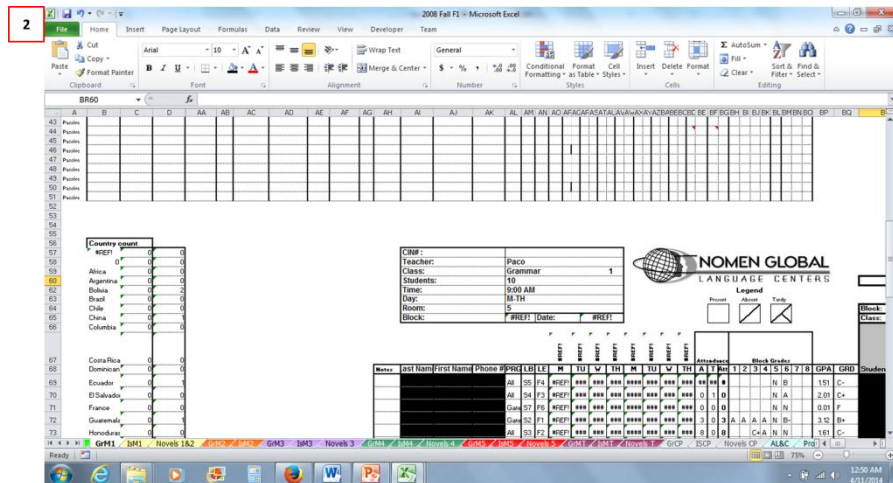
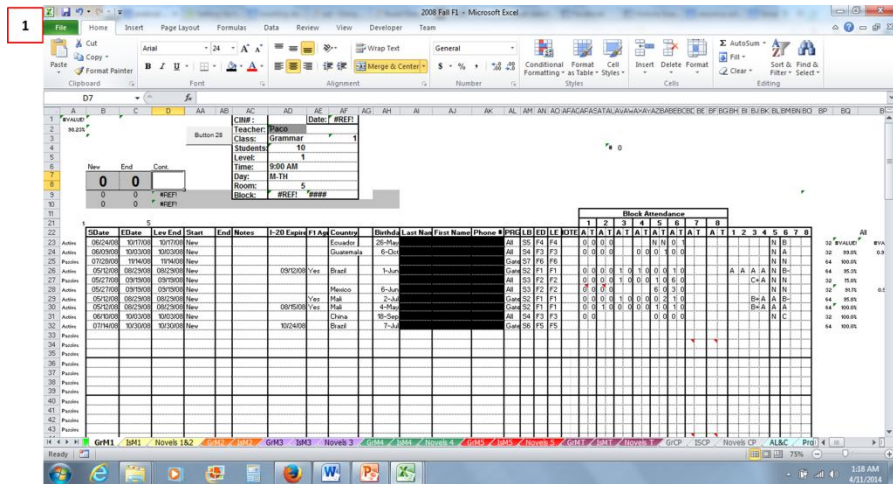
GRI %	GRI2 %	GRI3 %	GRI4 %	GRI5 %	calculate pass percentages and Numbers	GRI Pass	GRI2 Pass	GRI3 Pass	GRI4 Pass	GRI5 Pass
0.979403	0.788863	0.75418	0.767677	0.75418		134	148	148	148	148
0.564972	0.733735	0.733735	0.733735	0.733735		98	98	98	98	98
0.669386	0.733735	0.733735	0.733735	0.733735		108	108	108	108	108
0.47083	0.615786	0.615786	0.615786	0.615786		117	117	117	117	117

The *old* course-work performance data has quite a different story. Before 2013, Nomen Global had been using Excel files as a database for holding student course-work performance. I had to spend a couple of hours just *looking* at the information in order to understand how to even *begin* dealing with it: each workbook contained around 50 worksheets and it took a fair amount of time to figure out how they had even arranged their data amongst these worksheets. Finally, once I had identified the worksheets I needed, it took another little while to find where the correct information was located on the page.

More recently, I've found that although arrangement of the needed information on these worksheets is similar, it's not entirely consistent. The fact that the locations of information aren't perfectly consistent across these worksheets-of-interest means that looping through each of the sheets in exactly the same manner is not an option; therefore, I have found that dealing with this old set of data both has not and will not be as simple as dealing with the newer data has been.

*\*The next 3 screen captures are from the same worksheet in the old course-work performance dataset—I've included them to help convey how confusing this dataset has been. Please note the*

confusing arrangement of information (the third image contains the information I needed to access). Also, please note the small fraction of the total number of worksheets for this work book that are displayed on the bottom menu. Again, student information has been blacked out.



## 2.3

As previously stated, I have been helping Nomen Global justify their divergent teaching practices in teaching the English Language as a second language, so that they might secure substantial accreditation.

As stated in the executive summary, this was simpler for both the exam performance and the *new* course-work performance datasets. Because I manually imported both of these datasets into Excel, (I plan to try automating access to their web database in the future), I was able to arrange the layout in a manner that is well-suited for iterative processing (consult, again, the first two screen captures included).

For the *new* course-work dataset, I looped through the information and separately summed the number of students passing and the number of students failing. After the sums were obtained, I computed the percentage of students who passed by dividing the number of passing students by the sum of passing and failing students, and then displayed this information in the gray cells at the top of the sheet.

studentInfo_2013_Finalized - Copy - Microsoft Excel																				
File Home Insert Page Layout Formulas Data Review View Developer Team																				
Calibri 11 A A Wrap Text General Conditional Formatting as Table Styles Insert Delete Format AutoSum Fill Sort & Find & Filter Select																				
B I U Font Alignment Number Styles Cells Editing																				
A6																				
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	ER1 %:	0.979452		GR1 %:	0.748603		IS1 %:	0.75419							GR1 Pass	134		IS1 Pass	135	
2	ER2 %:	0.564972		GR2 %:	0.743719		IS2 %:	0.767677							GR2 Pass	148		IS2 Pass	152	
3	ER3 %:	0.79798		GR3 %:	0.732283		IS3 %:	0.666667							GR3 Pass	93		IS3 Pass	86	
4	ER4 %:	0.661765		GR4 %:	0.712418		IS4 %:	0.779221							GR4 Pass	109		IS4 Pass	120	
5	ER5 %:	0.47093		GR5 %:	0.615789		IS5 %:	0.541237							GR5 Pass	117		IS5 Pass	105	
6																				
7	10000172 LLTS	PRV		2013 W8	ER	ER4			817	0	0	F		0	0	4				Extensive Reading Level 4
8	10000172 LLTS	PRV		2013 W8	GR	GR4			829	0	0	F		0	0	4				Grammar Level 4

Similarly, for the exam performance dataset, I used looping to calculate the mean score, mean-improvement score, the variances of the scores, the variances of the improvements, and the number of students who improved. This information, similarly, was then displayed in gray cells at the top of the sheet.

combinedRetests_Finished - Copy - Microsoft Excel														
File Home Insert Page Layout Formulas Data Review View Developer Team														
Calibri 11 A A Wrap Text General Conditional Formatting as Table Cell Insert Delete Format AutoSum Fill Sort & Find & Filter Select														
Clipboard Font Alignment Merge & Center Number Styles Tables Cells Editing														
A2 It 1 Mean														
2	lt 1 Mean	361.2392857		lt 2 Mean	444.725		lt Improve Mean	102.125		lt Improve Perc	0.75		lt Improve Num	21
3	rt 1 Mean	460.1607143		rt 2 Mean	517.5071429		rt Improve Mean	76.00357143		rt Improve Perc	0.678571429		rt Improve Num	19
4														
5	gt 1 Var	34540.41534		gt 2 Var	35034.8072		gt Improve Var	17306.59206						
6	lt 1 Var	36853.82618		lt 2 Var	29450.22787		lt Improve Var	21766.75806						
7	rt 1 Var	26525.42099		rt 2 Var	13452.78587		rt Improve Var	20528.17739						
8														
9	Last Name	First Name	ID Number	Grammar Test			Listening Test	Reading Test						
10			5/13/2013	Grammar Test	506.7		Listening Test	437.5	586.8	Log 1-4				
11			8/13/2013	Grammar Test	573.9		Listening Test	613.8	653.5	Log 1-5				

This information has now been used by Nomen Global for generating reports, which indicated Nomen Global's 2013 students were performing just as well, if not better than, students who were taught with industry-standard methods (those reports are not part of my project).

My experience with Nomen Global's older data, however, has been much more difficult. As stated previously, the arrangement of information in the older dataset files is extremely confusing. I have written code that loops through the workbook and correctly selects the worksheets I need; furthermore, this code loops within each worksheet correctly and accesses the correct locations of information. Unfortunately, I am having difficult *now* because the information arranged on each of the sheets is not always accurate.

For instance, earlier in the semester, Dr. Allen helped solve one problem related to this: the sheet titled "GrM1" contained a hidden column, and so the command "activecell.offset(x,y)" was accessing the incorrect cell, even though (according to numerical counting) the "x" and "y" were correct. Dr. Allen resolved this problem with the command "'BQ" & ActiveCell.Row'.

At this point, the code for the *old* course-work performance dataset is almost completely finished; however, there are a couple of things I would like to address in the future: (1) I would like to establish a neater and more professional looking style for displaying the results, (2) I need to consult with Nomen Global and address a couple of questions regarding their information. Beyond those two goals, this dataset of *old* course-work performance includes *many* more files to be analyzed, and I am not certain if the code for "2008 Fall F1" will work properly (as it currently is) for all of the other files (for example, I may need to adjust certain cell indexes within the code for each of the other files). The display for the results is included in the following image.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	GrM1 Pass	7																			
2	GrM1 Fail	2																			
3																					
4																					
5	IsM1 Pass	0																			
6	IsM1 Fail	0																			
7																					
8																					
9																					
10	GrM2 Pass	4																			
11	GrM2 Fail	2																			
12																					
13	IsM2 Pass	5																			
14	IsM2 Fail	1																			
15																					
16	GrM3 Pass	6																			
17	GrM3 Fail	1																			
18																					
19	IsM3 Pass	8																			
20	IsM3 Fail	5																			
21																					
22																					
23																					
24																					
25																					
26																					
27																					
28																					
29																					
30																					

## 2.4

I learned several things from this project, but the *old* data files have taught me what I consider to be the most important. First, my experience with these files has taught me that *several* issues may be present pertaining to the information being dealt with (e.g., I experienced frustrations related to a

hidden column, cell values of " " (meaning, cells that look empty but aren't), and a confusing amount of information being held in a workbook and on the worksheets within that workbook).

In the future, as stated previously, I would like to try automating the retrieval of *new* information to be added to the *new* course-work performance dataset. Considering the arrangement of Nomen Global's online database, I believe this will take a lot of work.

I would also like to develop a neat way of displaying the information calculated for Nomen Global's *old* course-work performance dataset. Furthermore, I'd also like to extend the code I've written for the *old* dataset (currently "working" for the file "2008 Fall F1") to the rest of the files in that dataset (as previously stated, the arrangement of information in those files may not match perfectly with the arrangement in "2008 Fall F1," which would require some alterations to the code).

## 2.5

For this project, Dr. Allen helped me with the hidden column I encountered when dealing with the *old* dataset file "2008 Fall F1." This help from Dr. Allen was extremely significant, as I had tried everything I could think of in order to resolve the issue that was causing. Beyond this help received from Dr. Allen, however, I wrote the code completely on my own, but referred to examples on webforums as needed.