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.

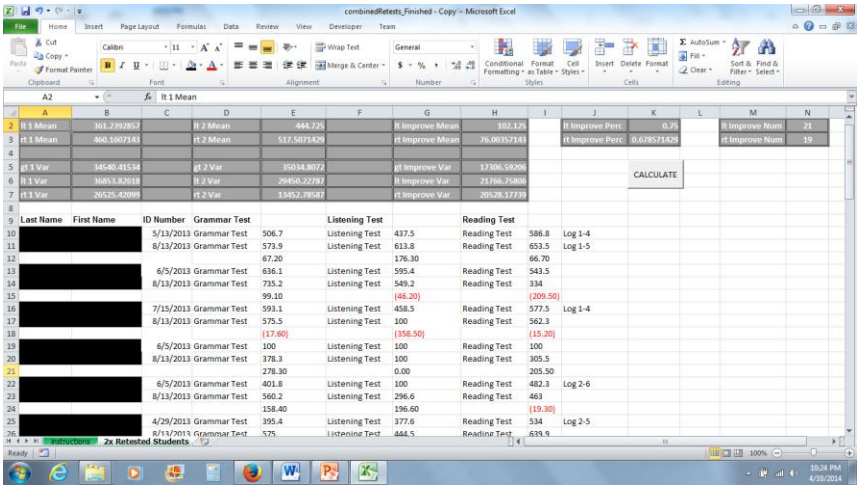
The company founder, a friend of mine, approached me in August 2013 requesting that I help with looking at the data Nomen Global had compiled on student performance. I have been working on a part-time basis to help Nomen Global justify its divergent program to the appropriate government agency. At the 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.

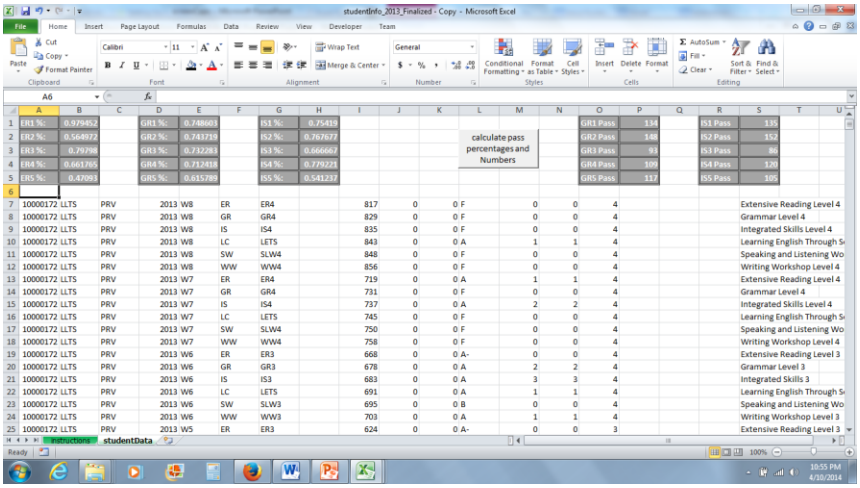
In my project, I have worked with three different sets of information for Nomen Global. Two of the data sets contain student performance in specific course work, and one set of data holds student exam performance (each datum compares a student's exam performance at the start of education and following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





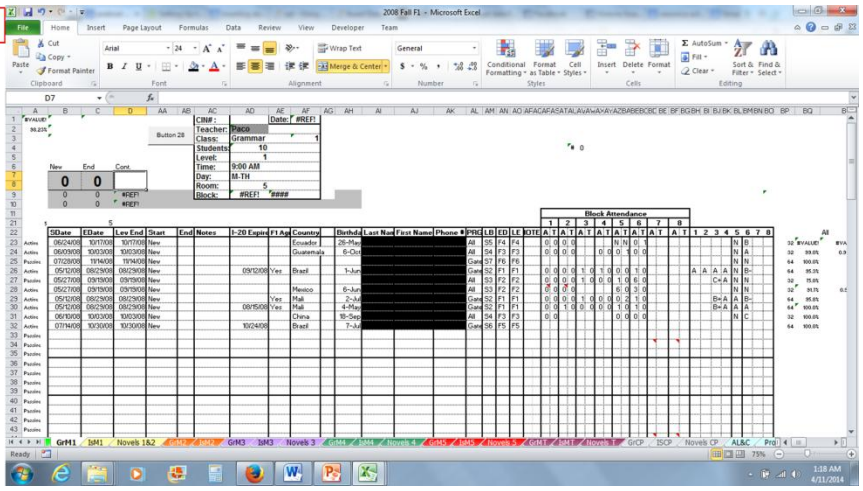
The *old* course-work performance data has quite a different story. Before 2013, Nomen Global had been using Excel files as a database for holding student course-work performance. I had to spend a couple of hours just *looking* at the information in order to understand how to even *begin* dealing with 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.

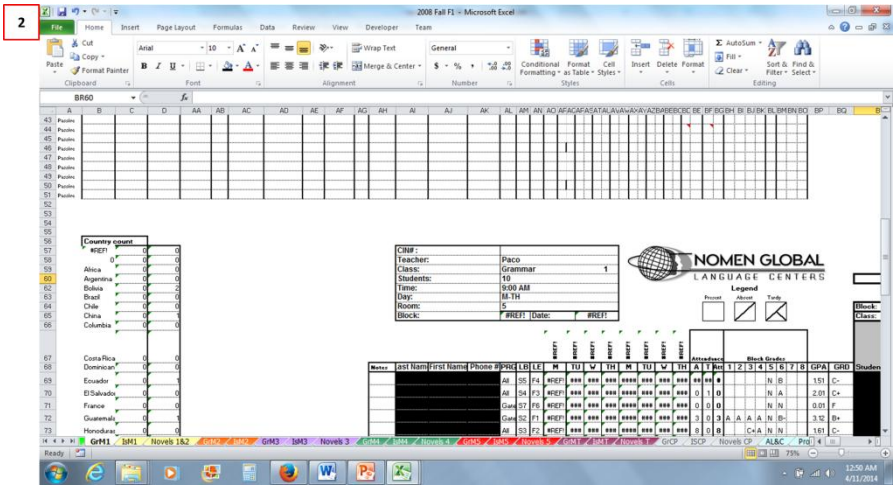
**The next 3 screen captures are from the same worksheet in the old course-work performance dataset—I've included them to help convey how confusing this dataset has been. Please note the*

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.

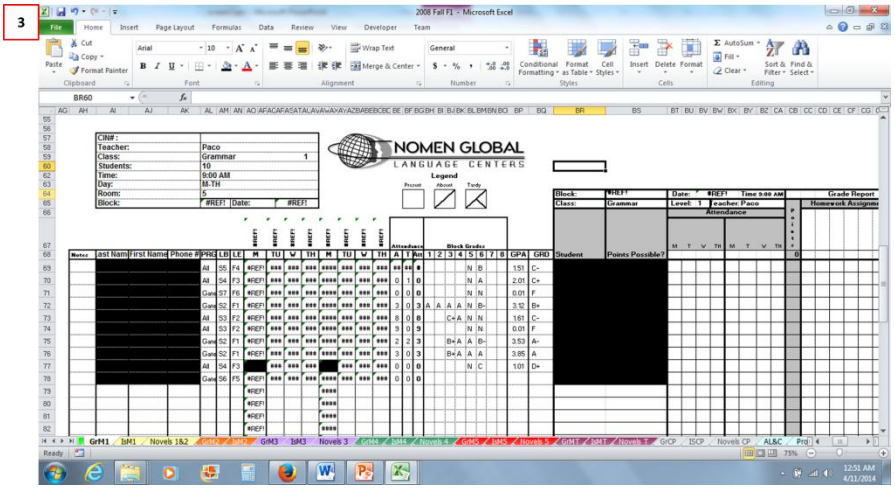
1



2



3

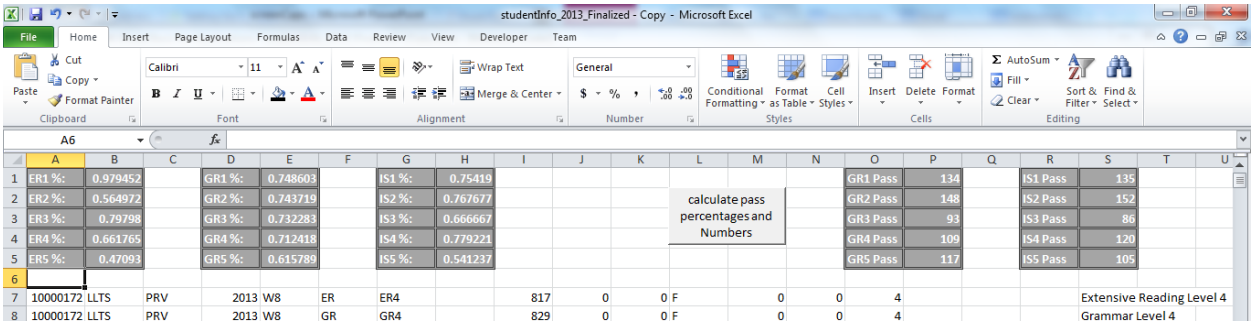


2.3

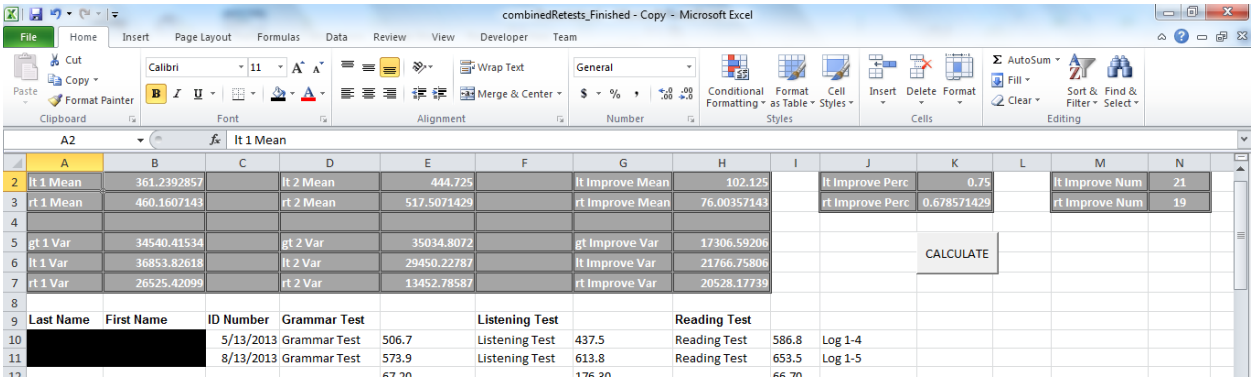
As previously stated, I have been helping Nomen Global justify their divergent teaching practices in teaching the English Language as a second language, so that they might secure substantial accreditation.

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.



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.

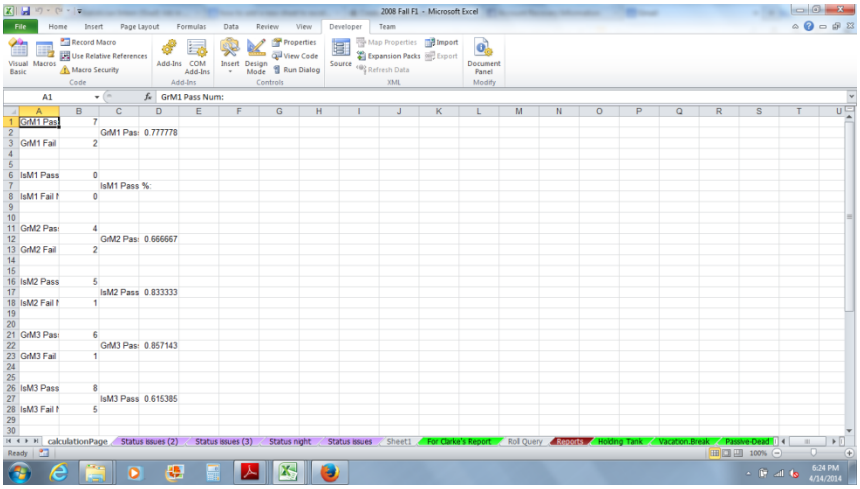


This information has now been used by Nomen Global for generating reports, which indicated Nomen Global's 2013 students were performing just as well, if not better than, students who were taught with industry-standard methods (those reports are not part of my project).

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.



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.