CS 1371 Lab 3: Using Microsoft Office Tools

What is Microsoft Excel?

Microsoft Excel is a spreadsheet application. Spreadsheets serve as efficient ways to organize data by grouping it into rows and columns. As a promising Tech student, you may never need to look at Excel after this lab. However, if one finds the need to organize numbers in a computer or perform administrative chores, Excel may soon become the best seven million lines of code ever written.

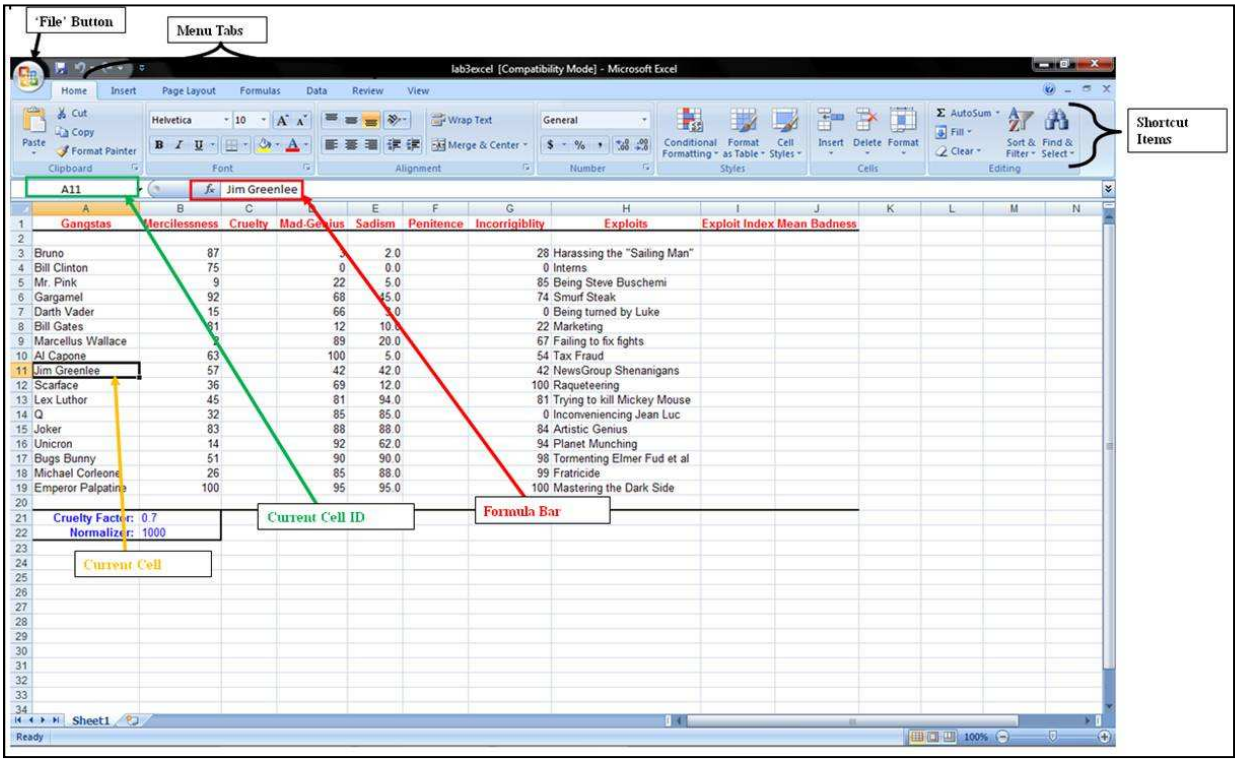
Also, it would be a good idea to recognize that when you are SAVING a 2007 document, you now have the option of saving it as an Excel 97-03 workbook OR as simply an Excel workbook. Since this lab requires that you learn to operate in Office 2007, **it is REQUIRED that your files be saved as ".xlsx" or ".docx". Failure to do so will result in a zero!**

Vocabulary

Cell -- the small boxes where data exists in a spreadsheet
Function -- with your extensive knowledge, no definition is necessary
PivotTable -- An interactive table that allows you to reorganize and summarize large amounts of data in a spreadsheet to obtain a desired report.
PivotChart -- An interactive chart that allows you to graphically present a large data set summarized in a Pivot Table.

Section 1

Retrieve the assignment (“lab3excel.xlsx”) from T-square under “Resources-> Lab 3” and open it using Excel.
(Note: in some browsers you will have to right-click on the link and select 'Save Target As'.)
You will see the following Spreadsheet:



Note that the first column in the spreadsheet is labeled A, and the first row is labeled 1. Similarly, each cell in the Excel spreadsheet has a unique identifier, much like the index of an element in an array. Like arrays in MATLAB, spreadsheets use a coordinate system to access elements. Unlike in MATLAB, letters are used to refer to columns, such that the upper- and left-most cell in the spreadsheet is called “A1”.

Also, notice that Excel 2007 looks much different than Excel 2003. There are no longer drop down menus at the top of the page – instead, there are tabs. This kind of interface is referred to as a “ribbon”. By clicking on each tab in the ribbon, you will bring up a new set of shortcut items. All of the standard options that used to be under the ‘File’ drop down menu (New, Save, Save As, Open, etc.) are now shown when you click on the Microsoft Office logo in the upper left-hand corner (shown in the image above).

Writing Formulas

Let us now work through some sample calculations that will aid in completing the rest of the lab. Let's figure out what Jim Greenlee's Exploit Index is.

The Exploit Index is defined as:

Exploit Index = Mad-Genius * Sadism

How can the Excel spreadsheet accomplish the same task? Each cell in a spreadsheet may contain either data or a formula. Move your cursor to where Jim Greenlee's Exploit Index should go (I11) and click there.

Now, rather than hard-coding the value as data, type in a formula to calculate Greenlee's exploit index. Type the following: =D11*E11 (Hit enter)

Note that Excel displays what you type in the formula bar--the row directly above the column guides.

Let's analyze the line that you just typed in. This line multiplies what is in cell D11, Jim's Mad-Genius, with the value from cell E11, Jim's Sadism, and displays the result in the spreadsheet cell I11. By clicking on I11, you can see the exploit index formula displayed in the formula bar. The value you should see in I11 is 1764.

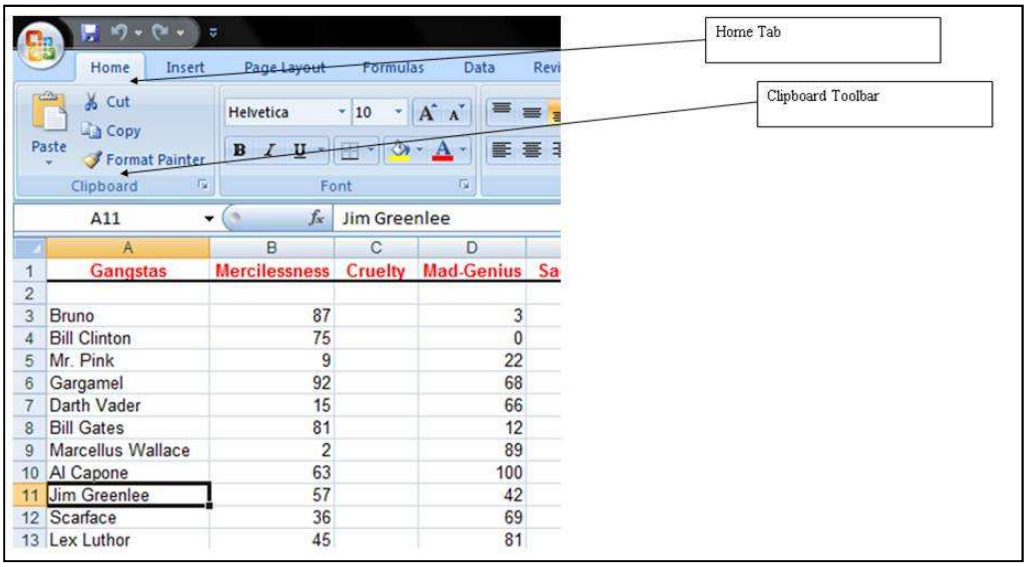
Copying Formulas

Now we should calculate the Exploit Index for everyone on our list. There are several ways to do this:

The Hard Way…

Repeat the procedure that you just applied on Jim Greenlee to everyone on your list. This solution takes a disproportionately long time to set up, but it will give you the correct answer. This method is, however, so inefficient that you may NOT use it in your submitted spreadsheet.

The better way:



	A	B	C	D	
1	Gangstas	Mercilessness	Cruelty	Mad-Genius	Sa
2					
3	Bruno	87		3	
4	Bill Clinton	75		0	
5	Mr. Pink	9		22	
6	Gargamel	92		68	
7	Darth Vader	15		66	
8	Bill Gates	81		12	
9	Marcellus Wallace	2		89	
10	Al Capone	63		100	
11	Jim Greenlee	57		42	
12	Scarface	36		69	
13	Lex Luthor	45		81	

Use the copy and paste commands in the ‘Clipboard’ area of the ‘Home’ tab (remember that each tab in the ribbon displays a different set of toolbars). When you copy a formula from one cell, the spreadsheet is smart enough to apply that formula to the other cells in a relative fashion. That means that if your

formula is in cell I11 and references cells D11 and E11, then when you copy it to I3, it will reference D3 and E3.

Step by Step:

1. Select the cell whose formula you need to copy (I11 in this case)
2. Select the copy option from the “Clipboard” tools. This copies the formula.
3. Now select all the cells to which you want to apply the formula (I3-I19).

   Hint: To select many cells at the same time, you can click and hold your mouse button on the first cell, and drag it and then release it on the last cell. You can also click the first cell, hold shift, and click the last cell.
4. Select the paste option from the “Clipboard” tools.
5. The calculated results should appear in all these cells.

Note: You can also copy from a cell by selecting it and then point your mouse to the bottom right corner of that cell. The pointer should become a bold cross. Now you simply have to click and drag over the cells you'd like to copy values to.

In the above calculation, we used a relative addressing scheme so that we could copy the formula easily but still make it apply to each individual row. Sometimes, however, we want a cell reference to stay constant. There are ways to use a value in just one particular cell as the source value. This is called absolute cell addressing.

Let's calculate the final value for the Exploit Index. This is accomplished by normalizing the previously calculated value for the Exploit Index:

Exploit Index = (Mad-Genius * Sadism) / Normalizer

The value for the normalizer is given in cell B22.

In Excel syntax, the formula for Jim Greenlee's Normalized Exploit Index would be given as:

=(D11*E11)/B22 <hit Enter>

The B22 means that you will be using the specific value in cell B22 for all your formulas, and this WILL NOT change when copying and pasting to other cells. This is called an “absolute” reference.

Precision and Formatting

For any number in Excel, you can set the precision for displaying it. Precision controls how many digits are displayed after the decimal point.

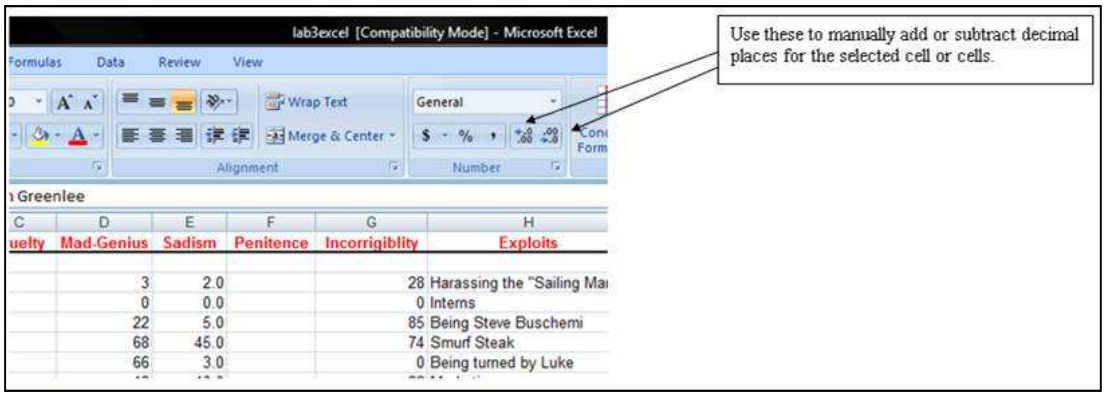
How to change a cell's formatting:

- Select the cell you want to adjust
- Right click on the cell

- From the pop-up menu that appears, select "Format Cells."

---OR---

(Refer to the following screenshot.)



Now format the appropriate values in your spreadsheet for Mad-Genius so that **exactly two**-decimal places are displayed on the screen. To format multiple cells concurrently, select all the cells that you need to format and then go to the format menu under the “Cell” region of the Home ribbon and select the “Format Cells” option.

Hint: Options are located under the “Number” tab of the Format Cells window.

There are other formatting options that are used to make spreadsheets more aesthetically pleasing. For example, you can change the text attributes of the labels, shade them, give them borders and generally change almost anything about the appearance of your spreadsheet. You may format your spreadsheet however you like, but do NOT change any of the values given!

Graphs

In Excel, a graph is defined as a diagram that displays a summary view of ~~the~~ data in a concise form. Below is a table that indicates situations where one type of graph can be more useful than another:

	Type of Graph	Useful When
	Bar Graph	Comparing several values measuring similar kinds of quantities.
	Pie Graph	Comparing percentages of a whole
	Line Graph	Looking at trends over equally spaced time intervals.
	Scatter Plot	Looking at trends (not necessarily at equally spaced time intervals)

Let us apply this to the spreadsheet we have been altering. Show a graph that displays each person's Exploit Index. Put the people on the X-axis and the Exploit Index on the Y-axis.

To plot the graph, go into the Insert tab and then select a Chart type from the chart toolbar.

You are responsible for figuring out how to draw the graph and which type of graph you choose to represent the data. Make sure that you put the graph on the sheet that you have been altering.

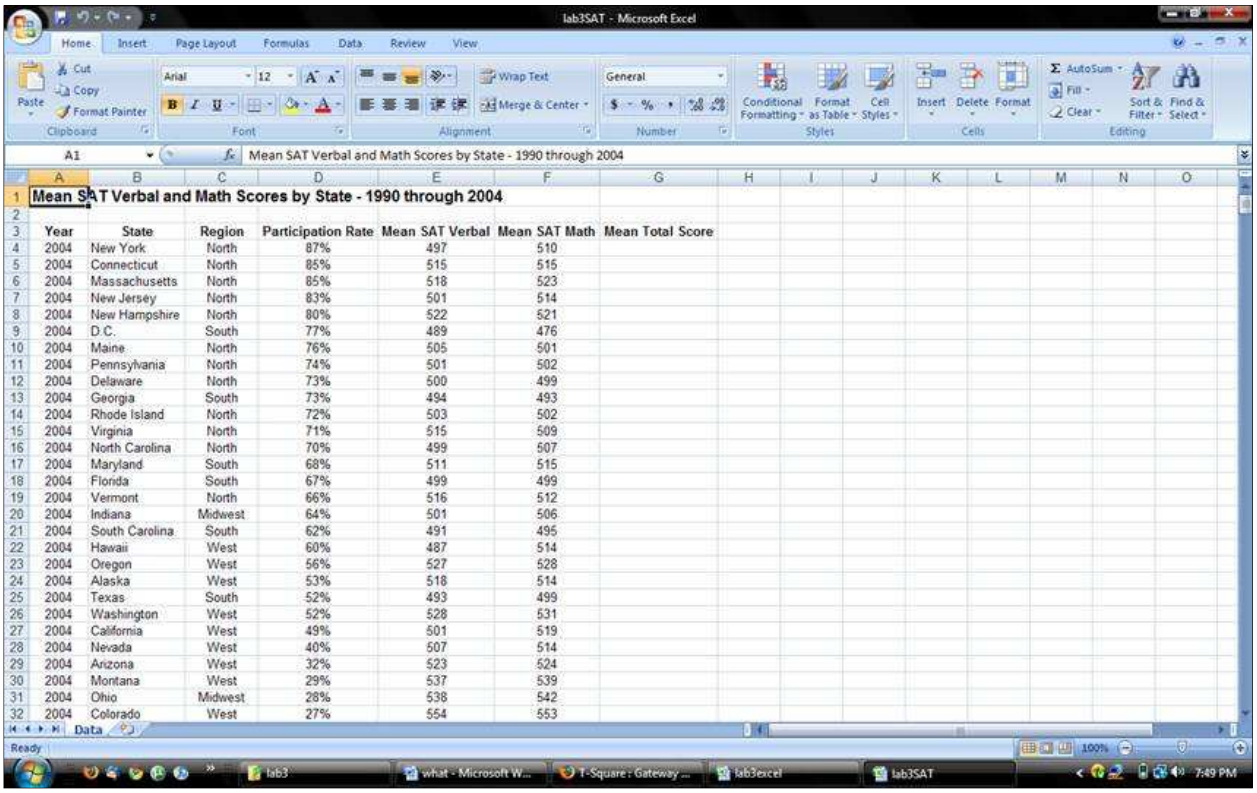
For the second part of the lab, please continue to section 2.

Section 2

Working with PivotTables

Retrieve the assignment ("lab3SAT.xlsx") from T-square under "Resources->Lab 3" and open it using Excel 2007.
(Note: in some browsers you will have to right-click on the link and select 'Save Target As'.)

You will see the following Spreadsheet:



Year	State	Region	Participation Rate	Mean SAT Verbal	Mean SAT Math	Mean Total Score
2004	New York	North	87%	497	510	
2004	Connecticut	North	85%	515	515	
2004	Massachusetts	North	85%	518	523	
2004	New Jersey	North	83%	501	514	
2004	New Hampshire	North	80%	522	521	
2004	D.C.	South	77%	489	476	
2004	Maine	North	76%	505	501	
2004	Pennsylvania	North	74%	501	502	
2004	Delaware	North	73%	500	499	
2004	Georgia	South	73%	494	493	
2004	Rhode Island	North	72%	503	502	
2004	Virginia	North	71%	515	509	
2004	North Carolina	North	70%	499	507	
2004	Maryland	South	68%	511	515	
2004	Florida	South	67%	499	499	
2004	Vermont	North	66%	516	512	
2004	Indiana	Midwest	64%	501	506	
2004	South Carolina	South	62%	491	495	
2004	Hawaii	West	60%	487	514	
2004	Oregon	West	56%	527	528	
2004	Alaska	West	53%	518	514	
2004	Texas	South	52%	493	499	
2004	Washington	West	52%	528	531	
2004	California	West	49%	501	519	
2004	Nevada	West	40%	507	514	
2004	Arizona	West	32%	523	524	
2004	Montana	West	29%	537	539	
2004	Ohio	Midwest	28%	538	542	
2004	Colorado	West	27%	554	553	

The data are organized into 6 columns containing information on Mean SAT Verbal and Math Scores by State for the last 15 years. Simply put, there is a lot of data here (765 rows by 6 columns). What if there was a way to organize all of this data into an easily viewable format so that we could actually make some use of it? We can do just that by putting all of these data into a PivotTable.

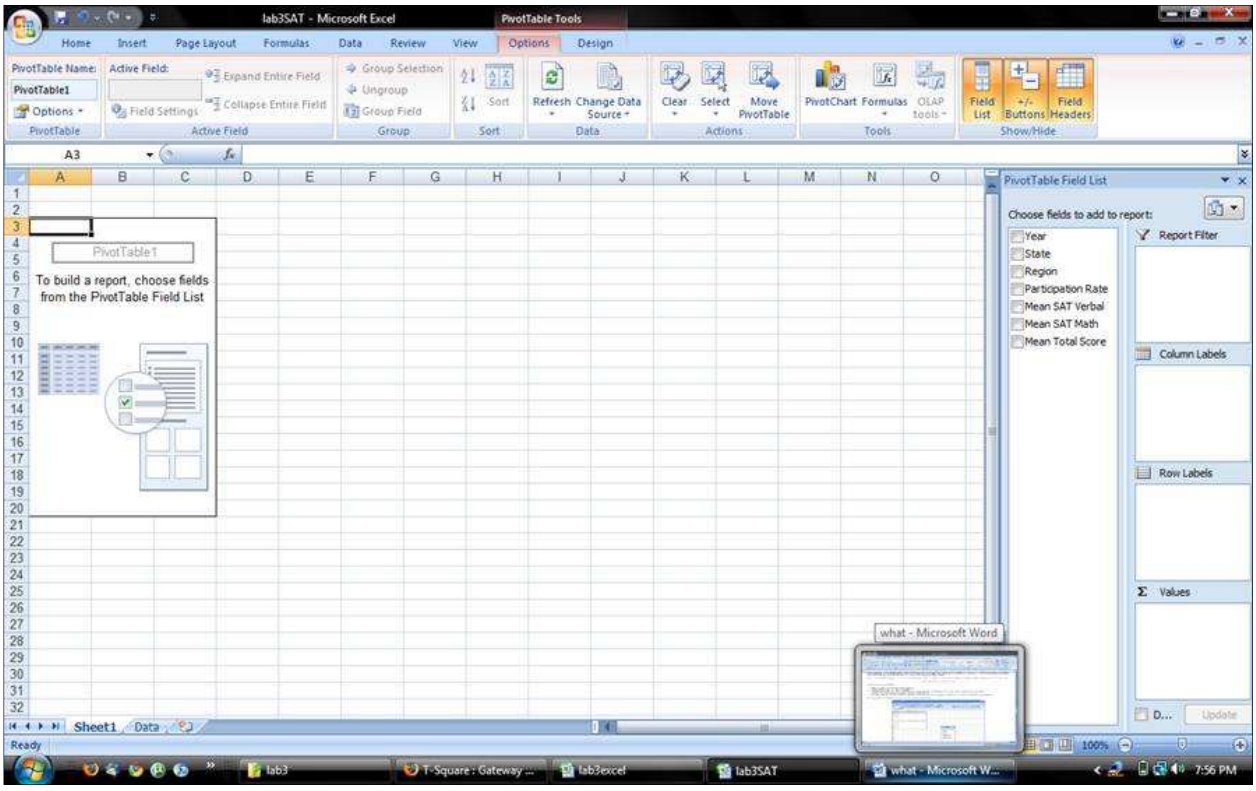
Since SAT scores are most often reported as a total score, the first thing we should do is calculate the Mean Total Score for each state and each year. We do this by writing a formula in each cell of Column G that computes Mean Total Score, defined as:

Mean Total Score = Mean SAT Verbal + Mean SAT Math

Next, we create our first PivotTable.

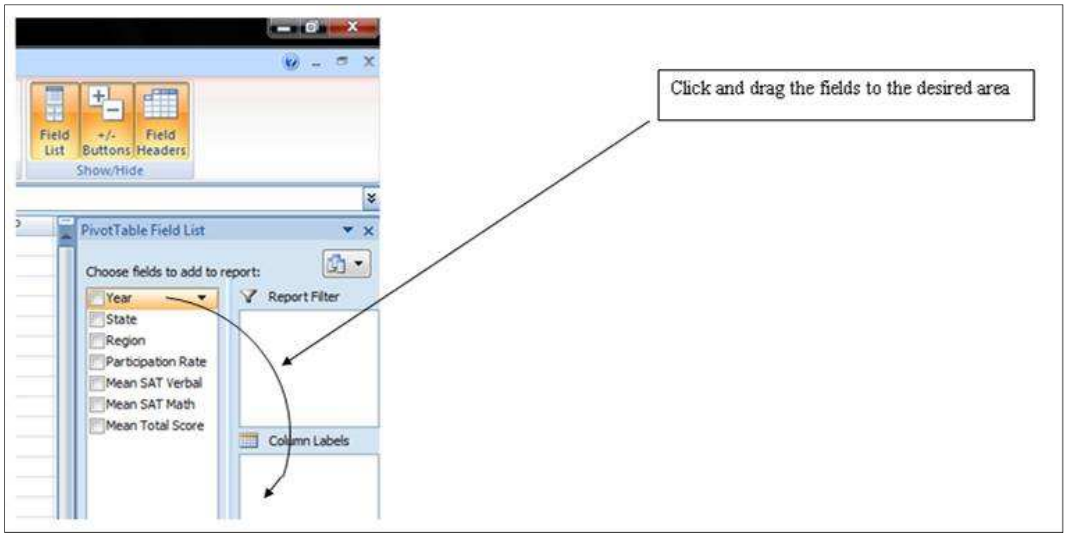
- Highlight all of the data you want to summarize, including any column titles.(“ Do not include the TITLE ON CELL A1 “Mean SAT Verbal and Math Scores by State – 1990 through 2004”)
- Click on the "Insert” Ribbon and select the PivotTable located on the ‘Tables’ toolbar. A new window named “Create PivotTable” will appear.
- Excel will ask you to choose the data that you want to analyze. These data should already be in the “Tables/Range” field. In the same window, Excel will ask you where you want to place the PivotTable report. Select "New worksheet" and click OK.

A new worksheet will have been added, and your file will now look like this:



You will notice that there are four areas to which you can add fields to the PivotTable. We could conceivably analyze the data in any number of ways, but for this assignment we'll be looking at the Average of Mean Total Score. To add a field to an area in the PivotTable, simply select it from the PivotTable Field List and drag it to the desired area.

- Select the "Region" field and add it to the "Report Filter" area.
- Select the "Year" field and add it to the "Column Labels" area.
- Select the "State" field and add it to the "Row Labels" area.
- Select the "Mean Total Score" field and add it to the "Values" area.



You will notice that cell A3 is colored differently than most of the cells, and it says "Sum of Mean Total Score". This tells us how we are summarizing the data. Recall from earlier that we said we wanted to display our data as the Average of Mean Total Score.

To change this, go over to the right side of the screen and click on the “Sum of Mean Total Score” bar in the “Values” section (the one that appeared there when you dragged it in). Now select the “Value Field Settings” from the drop down menu, then change the highlighted option from “Sum” to “Average” and then click OK.

Your summary columns should now contain the average score for each state and each year. Some of these will have too many decimal places – we want reduce it to one decimal place. Do this by highlighting the last row (highlight the numbers only) and change the decimal places to 1. Do this for the last column as well.

You will now see a PivotTable containing a summarized version of the original data, with averages for each year and each state. The data are sorted alphabetically by state. This is useful, but what if we wanted to see what state has the highest Average SAT Score over the 15-year period?

Click on the options TAB and from the ‘Sort toolbar select Z-> A ICON (This will sort in descending order).

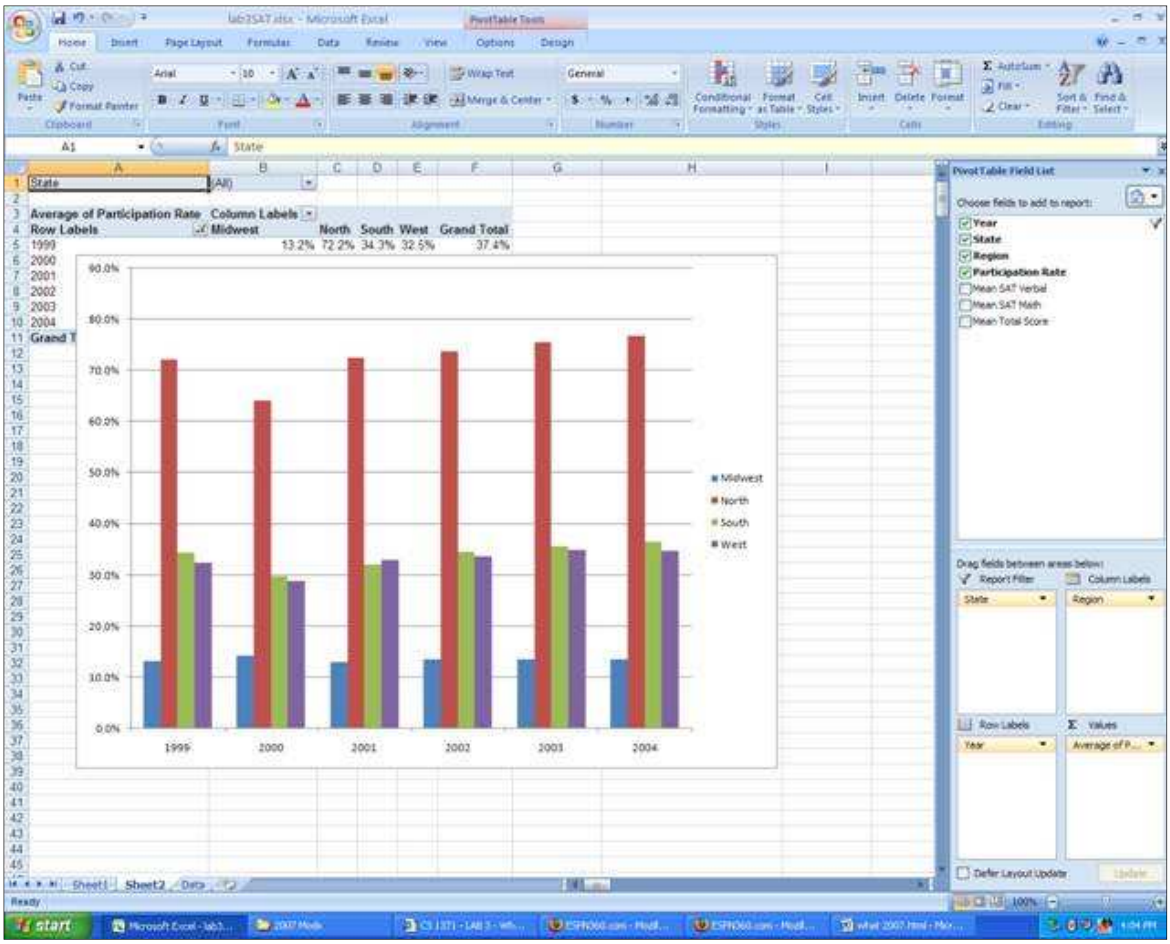
Now our data are sorted, and we can easily see that Iowa has the highest Average Total Score over the 15-year period. Georgia is third from the bottom.

Working with PivotCharts

We've successfully created a PivotTable in Excel. Now we would like to create a graphical representation of our data. We can accomplish this using a PivotChart. We can just easily create a

PivotChart from the PivotTable we have just created but for the purpose of this LAB we will create a new PivotTable from which we will create our PivotChart.

- Go back to the worksheet containing the original data, and highlight all of the data including the column titles (Again, do not include the TITLE ON CELL A1 'Mean SAT Verbal and Math Scores by State - 1990 through 2004')
- Click on the 'Insert' TAB on the upper part of the screen and select the PivotChart located on the 'Tables' toolbar. To see the PivotChart option, click on the small arrow right below the PivotTable ICON (dropdown menu). A new window named 'Create PivotTable with PivotChart'
- Excel will ask you to choose the data that you want to analyze. These data should already be in the 'Table/Range' field. In the same window, Excel will ask you where you want to place the PivotTable and Pivot Chart report. Select "New worksheet" and click OK.



As before, there are four areas to which you can add fields to the PivotChart. In creating our PivotChart we'll be looking at the Average Participation Rate by Region. Just as we did before, to add a field to an area in the PivotChart we select it from the PivotTable Field List and drag it to the desired area.

- Select the "Region" field and add it to the "Legend Fields" area.
- Select the "State" field and add it to the "Report Filter" area.

- Select the "Year" field and add it to the "Axis Fields" area.
- Select the "Participation Rate" field and add it to the "Values" area.

This time the data are summarized by "Count of Participation Rate", and data are displayed for the years 1990-2004. However, we want to view the data by Average Participation rate. Also, we don't have a complete data set, so we will need to change the years for which we display the data.

- Do this the same way you changed the data in the pivot chart (see above) Also, in the same box where you change the data from “Count” to “Average”, click "Number" and select the "Percentage" category to display the numbers as percentages.
- To change the years displayed, click on the “Year” bar in the PivotChart Filter Pane, and select only the years 1999 - 2004.

We're almost finished, but the data are still a little hard to see. To fix this, we can change the type of chart we've created.

Do this the same way you changed the data in the PivotTable before. Also, in the same box where you change the data from “Count” to “Average”, click "Number Format" and select the "Percentage" (one decimal place) category to display the numbers as percentages.

To change the years displayed, click on the arrow next to the “Year” field in the PivotChart Filter Pane, and select only the years 1999 - 2004.

We're almost finished, but the data are still a little hard to see. To fix this, we can change the type of chart we've created.

- Click on the “Design” TAB and on the “Type” toolbar select “Change Chart Type”.
- Under the column chart type select the "Clustered Column" chart sub-type (this is the one in the upper left corner).

Note: your chart may already be a clustered bar graph. If so, you don’t need to worry about this step.

Now we've finished creating our chart, and we can see that the highest SAT Participation Rate occurs in the Northern states. If we wanted to compare just two regions rather than all four (say, the North and the South), we could click on the arrow next to the "Region" field on the PivotChart Filter pane and select just the North and South checkboxes.

The Assignment (Part 1)

1. Get the Excel file you will be editing (you should have done this already if you read the lab).

2. Calculate the Normalized Exploit Index for all the people listed (you should have done this already if you read the lab).

3. Fill in the Mad-Genius column to 2 decimal precision (you should have done this already if you read the lab).

4. Plot a graph of the Normalized Exploit Index vs. Person's Name (you should have done this already if you read the lab).

5. Fill in the Cruelty column. The formula for Cruelty is:

   Cruelty = Cruelty Factor * Mercilessness

   Hint: The absolute reference information will be useful here.

6. Use Excel to fill in the Penitence column. Penitence is defined as:

   Penitence = 1 / Incorrigibility

   You will note, however, that there are a couple of values in the incorrigibility column that have a zero numeric value. Dividing anything by zero is a bad idea. In MATLAB you would get infinity, but in Excel you would simply get the error message “#DIV/0!”. In order to avoid this pitfall, simply leave these cells blank. All others that are divisible must be evaluated.

7. Calculate the Mean Badness of each individual listed in the spreadsheet.

   Mean Badness = (Mercilessness + Cruelty + Mad-Genius + Incorrigibility) / 4
   It is left to you to translate this equation into a valid Excel formula.

8. After you have calculated the Mean Badness for each and every villain, use Excel's built-in sort feature to sort the villains in ASCENDING order based on their Mean Badness.

9. Create a graph of each villain (X-axis) vs. Mean Badness (Y-axis). Make sure it is on the same spreadsheet as the data.

10. Save this file as "lab3excel.xlsx".

11. The second file you create should contain 3 worksheets and 1 chart. The original data sheet (the one we gave to you) should be in a worksheet called "Data". The first PivotTable you created should be in a worksheet called "Sheet1". The PivotChart you created should be in a worksheet called "Chart1". There is also another worksheet, but the name of this sheet doesn't have to be

anything specific. If you don't name the worksheets/charts EXACTLY as we have asked, you will lose points on your lab assignment.

12. Call the first excel file "lab3excel.xlsx" and the second file "lab3SAT.xlsx".

DON'T FORGET THAT THERE IS A SECOND PART TO THIS LAB!

Important Notes:

Look at direction #11 closely….THIS IS IMPORTANT! You will lose points if your worksheets are not named "Data", "Sheet1", and "Chart1".

Refer to "Lab 3 Part 2.pdf" for instructions regarding the second part of the lab.