Lecture Note #16: PivotTables Part #2

BUSI 201: Business Data Analysis

Fall 2023

Topic 1. PivotTables: Refreshing Data

One of the reasons that we may favor the use of functions over PivotTables is that functions will automatically update its values when information is edited in the original data. By default, PivotTables do not automatically update, as it does not draw directly from the original data. Instead, PivotTables refer to cached data, so we must manually update the PivotTable.[1]

To learn how to refresh the data on our PivotTables, navigate to worksheet `PIVOT-06` of the workbook `BUSI201-LEC16-Workbook.xlsx`. This worksheet contains a randomly generated gradebook for some class. Suppose that we are interested in analyzing the students' performance data based on their majors and class. Following the steps illustrated in Figure 1.
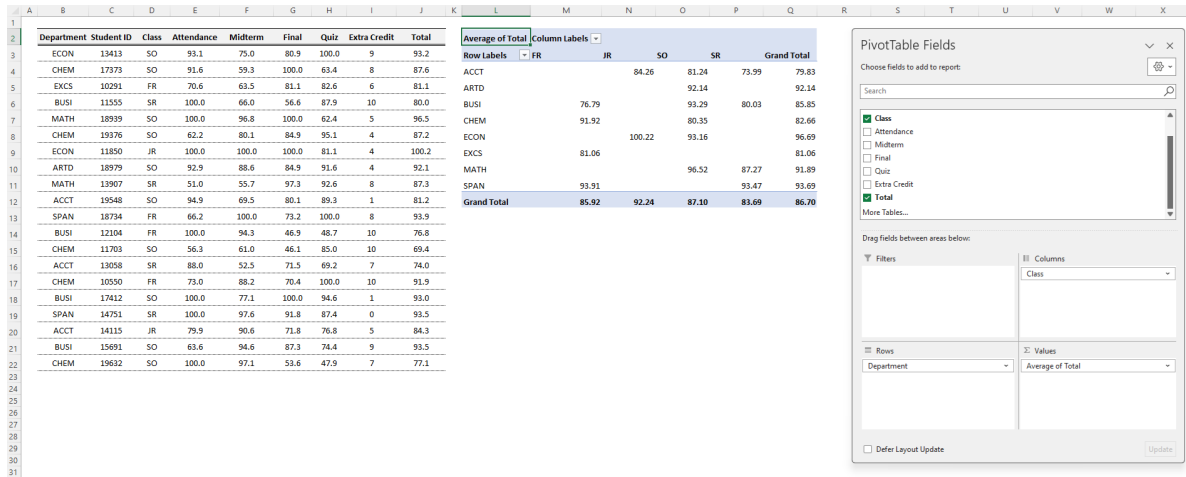

Figure 1: `PIVOT-06`

Suppose that there were some last minute changes where some students received extra credit after their final exams. Lets imagine that the student with 0 extra credit in row 19 has completed a task that earned them 10 points extra credit. There is only one senior majoring in Spanish in this dataset, so ideally speaking, the PivotTable should reflect this change.

Figure 2: PivotTable Not Synced

See Figure 2. Note that even after updating the original data, the PivotTable did not update accordingly. If we want the PivotTable data to be updated, we can manually initiate the process. While there are many ways to initiate this process, and we will be covering three methods.
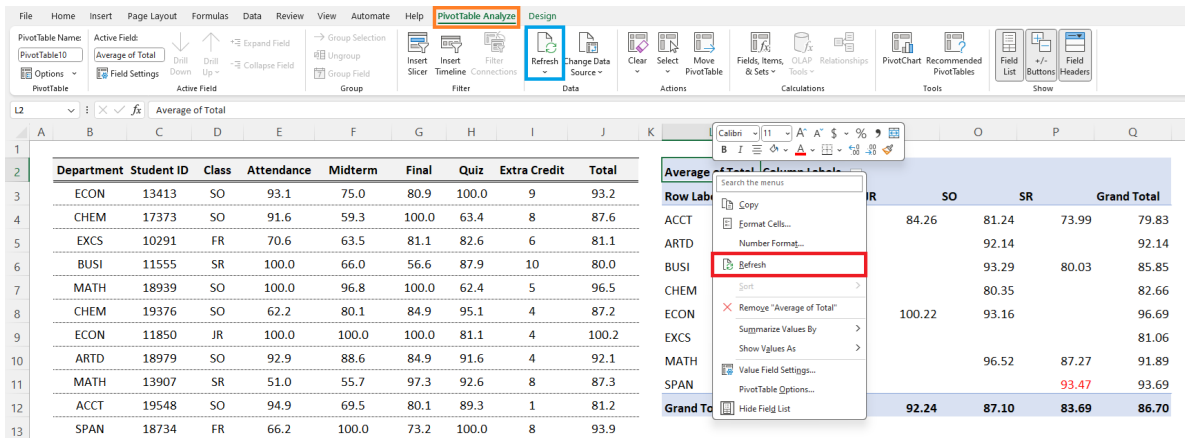

Figure 3: PivotTable Refreshing

First, select any cell on the PivotTable to pull up menu items on the ribbon on the top side of the Excel window. Then, select `PivotTable Analyze`, and then select `Refresh`. This should refresh the PivotTable so that the new extra credit score is reflected in the table. The second way of refreshing the PivotTable is also quite simple. Select any cell on the PivotTable, right click once, and select `Refresh`. The last method is using the hotkey alt + F5 after selecting any cell in the PivotTable.

PivotTable Settings: Autofit Columns

Actually refreshing the table, you may have noticed that the width of each column changed to fit the data in each cell. Suppose you want to keep the width of each cell/column constant even after you refresh the PivotTable. We can change the settings so that they remain the same width after refreshing the data.
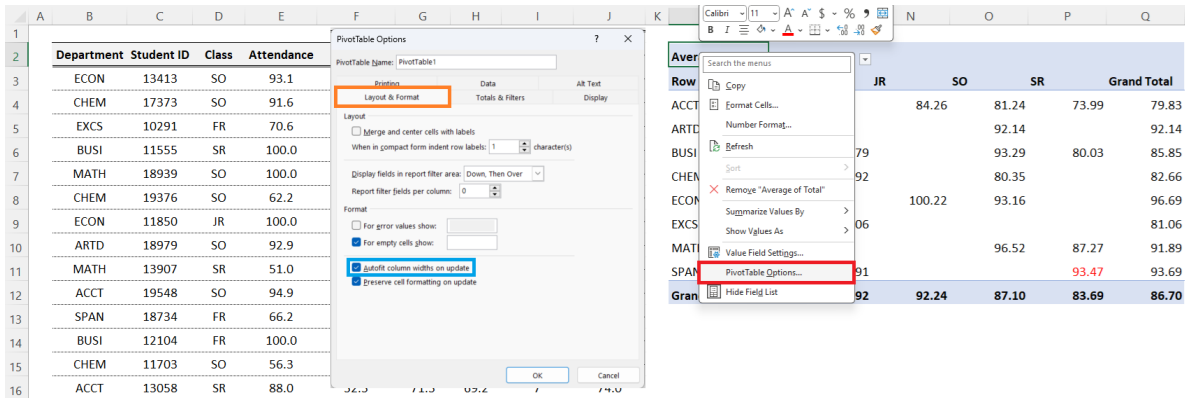

Figure 4: PivotTable Options: Autofit

Select any cell on the PivotTable, right click, and select `PivotTable Options`. Navigating to the `Layout&Format` tab in the new pop-up window, and deselect `Autofit column widths on update`. Following this process, the column widths will remain constant each time you update the PivotTable.

PivotTable Settings: Manually Sorting Columns

One thing that bothers me is that the `Class` is not sorted properly. We as humans understand that the correct order should be FR>SO>JR>SR, but Excel does not. We can manually sort the columns by following a few steps. First move your mouse cursor to the upper border of cell `N3` so that the mouse cursor changes to a downward pointing arrow. Left click once, and then move the cursor slightly upward until the cursor changes to a "move" cursor (looks like a + sign with arrows pointing outward). Then, you can click and drag the entire column.
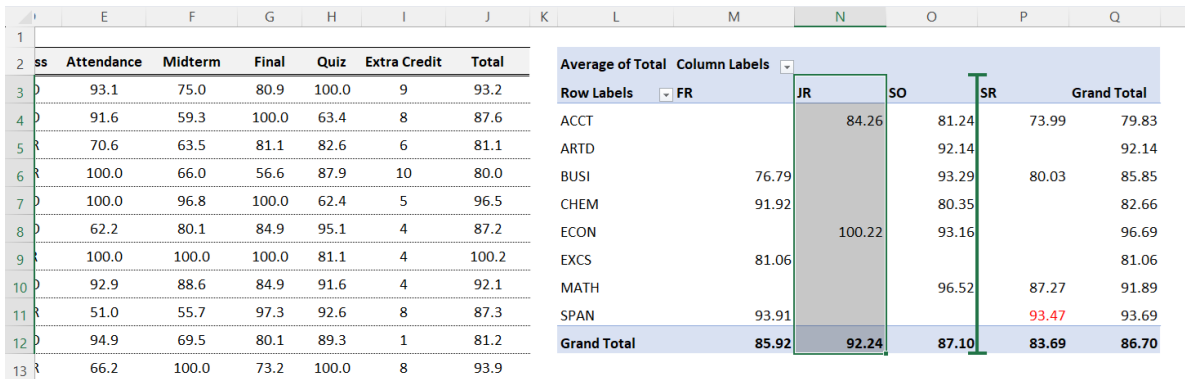

Figure 5: Manually Sorting PivotTable Columns

Topic 2. PivotTables: Declaring "Tables"

Recall how we started off building PivotTables. We selected the range that contained the data, and then selected the location where the PivotTable should show up, and we set up the PivotTable using the new menu that pops up. One problem with this approach occurs when we want to add entries to the original data.

Suppose you add a new student to the list in worksheet `PIVOT-06`, where you already have a Pivot-Table, and then try to update the PivotTable. This new line of data will not be automatically added to the PivotTable, since the range that we selected when setting up the PivotTable does not include the final row. We will learn how to declare tables so that we can fix this problem.
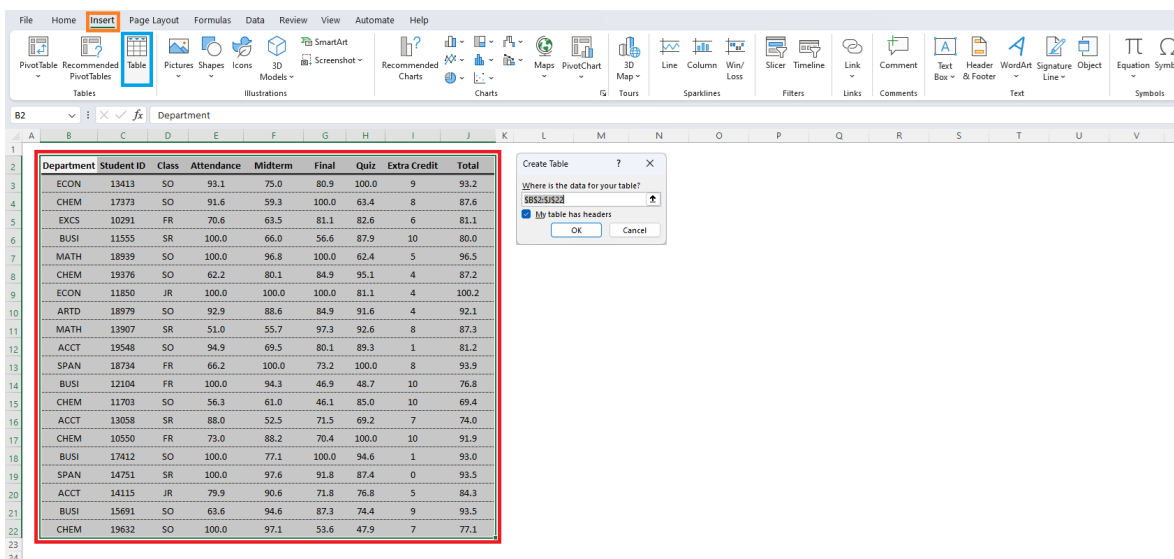
Figure 6: Declaring the Data as a Table

Navigate to the worksheet `PIVOT-07`, which is a duplicate of the previous worksheet. We will be selecting the dataset we will be working with, and declare it as a table. Select the entire dataset including the variable names at the top row as shown in the red box of Figure 6. Then, navigate to `Insert`, and then select `Table`. There are some conditions that should be satisfied for this process to work out:

1. No merged cells should be included in the selected dataset.

2. There can be no duplicate variable names in the first row.

3. All variable names must be included in a single row.
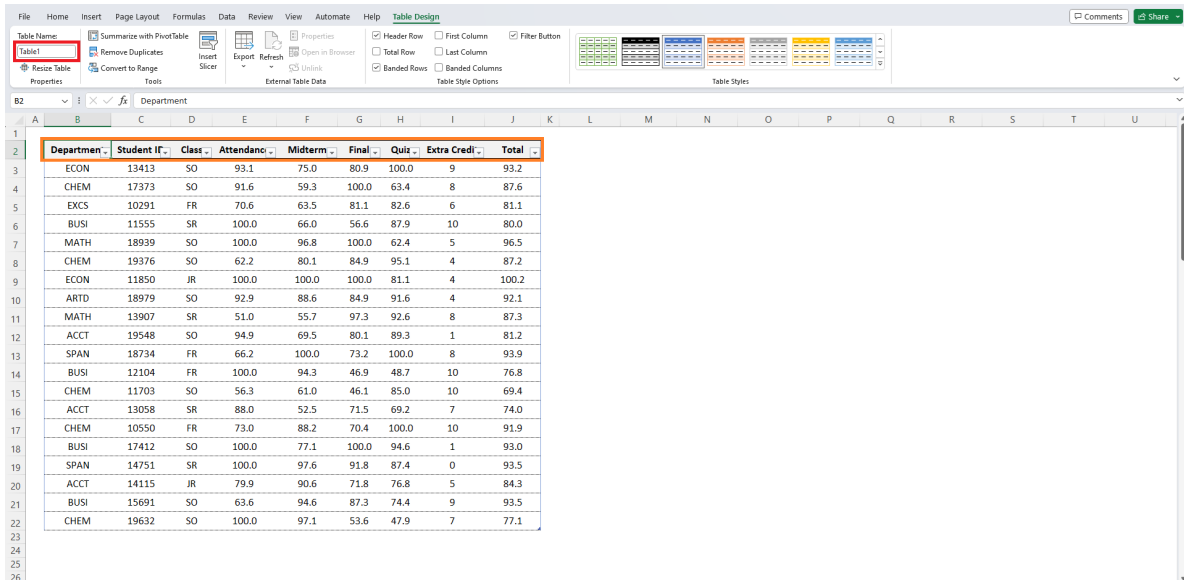
Figure 7: Declaring the Data as a Table

We can name the table by changing the terms in the red box in Figure 7. We should keep track of this name, as we will be referencing this when we create the PivotTable. If the dataset is correctly declared as a table, you will see filters being automatically added as shown in the orange box in Figure 7.
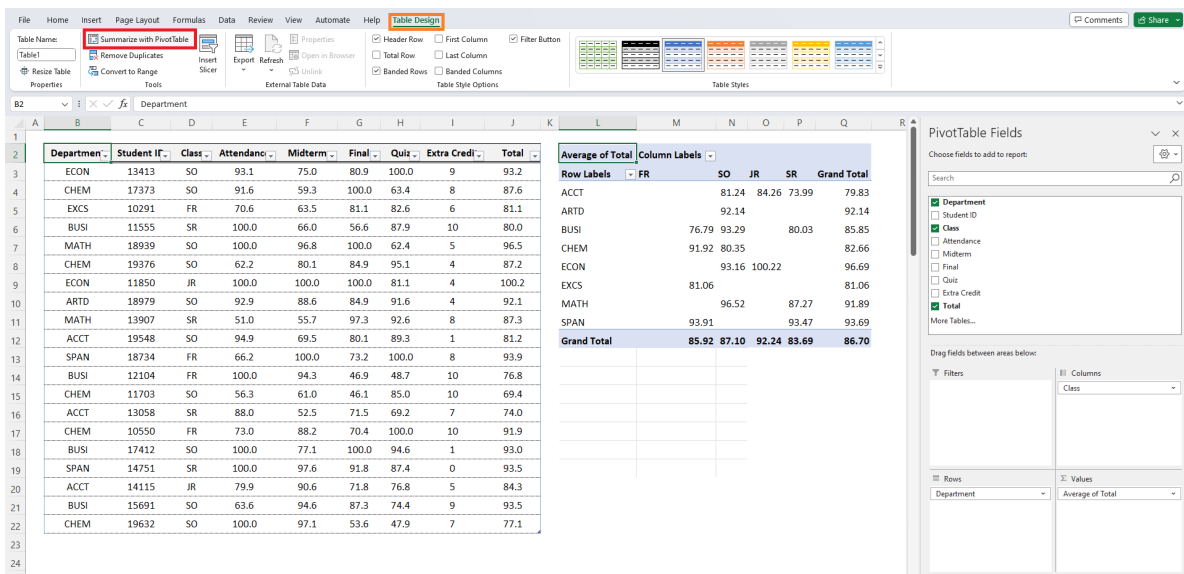
Figure 8: Generating PivotTables with a Table

Select the table that we just declared, and then select `Table Design` in the ribbon menu. Then, select `Summarize with PivotTable`. Following all procedures to generate the PivotTable, we can reach the state illustrated in Figure 8. Now lets see what happens when we add a new entry.

Figure 9: Adding Data to a Table

Suppose that we add a student below the very last row of the data. See the entry generated in the red box in Figure 9. Notice that the PivotTable does not automatically update, since the cell that should now have data is empty as shown in the orange box in Figure 9. However, you may have noticed that the table itself has been expanded to include the newest entry, as evidenced by the blue outline expanding to surround the new row.

Figure 10: Updated PivotTable

Refresh the PivotTable using any of the three methods we learned in the previous section. You should notice that the newly added row is now added in the PivotTable, as the entry shows up in the orange box in Figure 10.

Topic 3. PivotTables: Formatting

While there are many formatting options available for PivotTables, we will be covering some of the basic ones that will be useful in real-world scenarios in this chapter. Navigate to worksheet `PIVOT-08` for an expanded version of the gradebook from the previous worksheets.

Figure 11: Default PivotTable

Number Formats

For most purposes, we do not need any more than 2 decimal points. We can change the formatting of the cells in the PivotTable to make the table more readable. Simply select any cell on the PivotTable, right click, and select `Value Field Settings`. Then, click on `Number Format` to call up the formatting window in the orange box. You can change the format of the numbers in this window, and apply the effect to the entire PivotTable.

Figure 12: Number Formats

7

Report Layout

The default layout of the table is `Compact`, which minimizes the number of columns by stacking variables in a single column. Although it has its advantages, it is often better to experiment with other layouts. Lets edit the PivotTable to be a bit more complex for our future analysis.

Figure 13: Layouts

The default layout of PivotTables is the `Compact` layout, which minimizes the number of columns by placing all row variables in one column. This format takes up less space, but it is not optimal for data analysis purposes. We want each row variable to populate its own column, since we may have to set up conditions based on these variables. Select `Design`, and then `Report Layout`, and select `Show in Tabular Form`.

Figure 14: Tabular Layout

Other Options

The readers are encouraged to try out other formatting options in the `Design` tab. There are options that allow you to enable/disable the subtotals for each category in the red box. You can also generate or remove grand totals which take up the very last row/column of each PivotTable in the blue box.

Figure 15: Other Options

The options in the orange box allows you to add an empty row between categories, and the options in the green box automatically shades alternating rows/columns which assist on reading tables. To change the colors or borders of the PivotTable, you should look into the options in the purple box.