Lecture Note #17: PivotTables Part #3

BUSI 201: Business Data Analysis

Fall 2023

Topic 1. PivotTables: Grouping

When we have a wide range of variables taking up the row or columns, we may gain more insight about the data when we gather them into larger groups. For instance, you may want to analyze individuals by groups of age, daily sales data in terms of weeks, etc. Navigate to worksheet PIVOT-DATE in the workbook BUSI201-LEC17-Workbook.xlsx. Replicate the table below using PivotTables:

		1/1/2023	1/2/2023	1/3/2023	1/4/2023	1/5/2023	1/6/2023
Branch 1	Item 1						
	Item 2						
	Item 3						
Branch 2	Item 1						
	Item 2						
	Item 3						

Filling each empty cell with the number of each items sold in each office by date, we get the table in Figure 1. The problem with this table is that it is quite "wide," where the total number of columns add up to 34. So, we may want to aggregate sales data by week to reduce the number of columns in our table to have the table fit in one reasonable page.
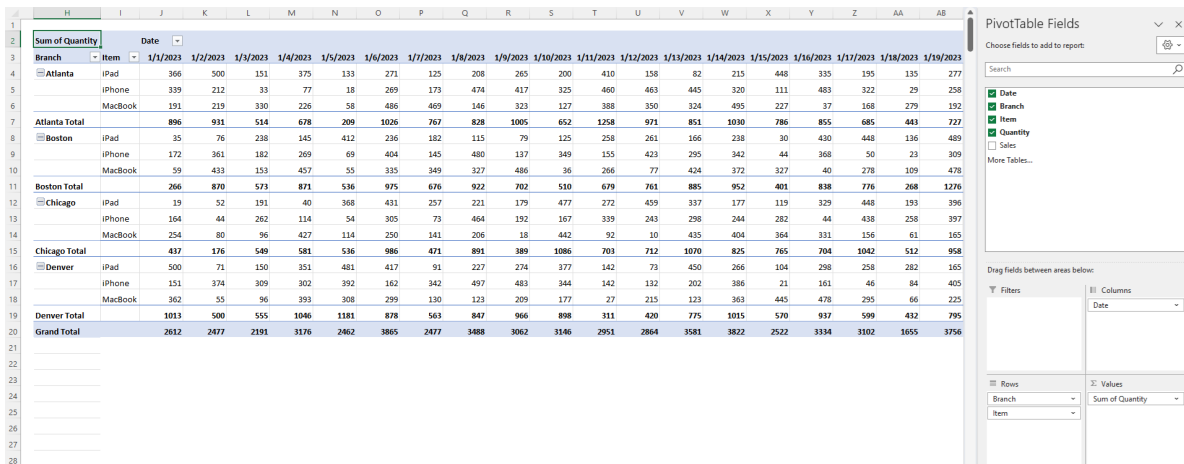

Figure 1: PIVOT-DATE

Grouping by Date

To group dates into weeks, first click on any of the dates in the PivotTable, right click, and then select `Group`. In the `Grouping` window that pops up, select `Days`, and then change the `Number of days` to 7 to group days into weeks.

Figure 2: Grouping by Dates

The PivotTable with dates group into 7 days each shows up in Figure 3. You may want to change the number of days so that each column represents a similar number of days, or adjust the beginning and end dates before grouping the variables so that the table is "regular."

Figure 3: Grouped by Dates

Grouping by Numbers

We can also group the variables as long as it is numerical. Navigate to worksheet `PIVOT-YEARS`, which houses real world data on government expenditure as a proportion of GDP. One of the variables is `Year`, but note that it is not automatically recognized as dates. Generate a PivotTable with years being the row variables, Country being the column variables, and each value entry to be the average of G/GDP.
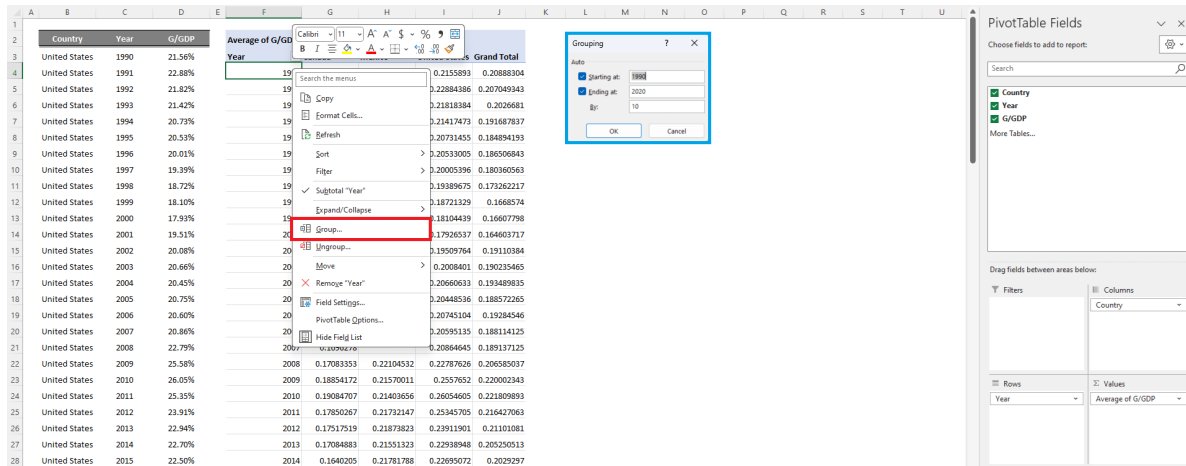

Figure 4: Grouping by Numbers

Right click any year in the PivotTable, and select `Group`. A slightly different grouping window will pop up in the blue box, as Excel does not recognize the years as a date variable. Figure 5 shows us what the PivotTable should look like when we group the data by 5 years.
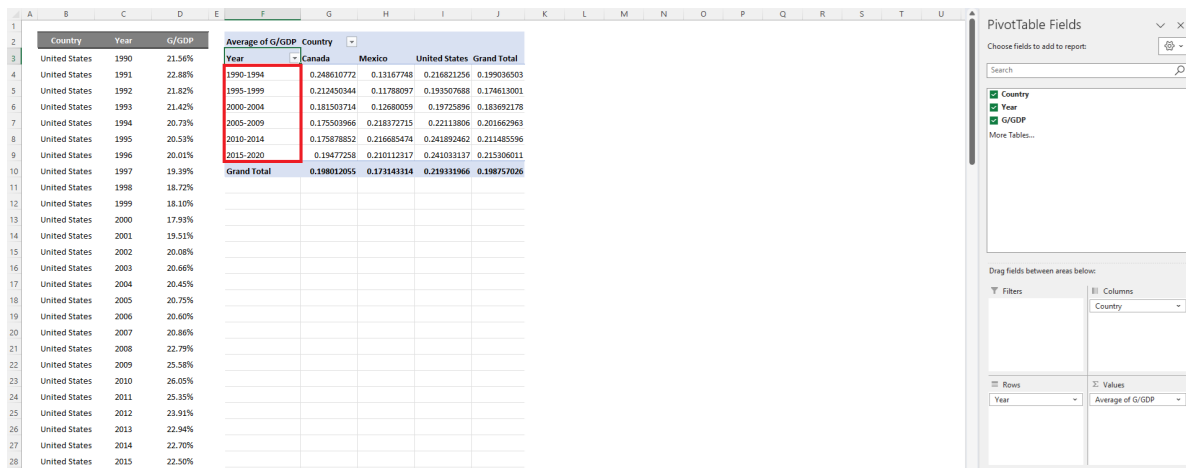
Figure 5: Grouped by Numbers

Topic 2. PivotTables: Filters

Up to this point, we have been primarily interested in three "boxes" when we were creating Pivot-Tables; the Rows, Columns, and Values. However, the remaining fourth box, Filters, can also be quite useful. Adding a filter variable will create a filter above the PivotTable which we can use to filter the data. Let us try out this feature to understand what this filter does for us. Navigate to `PIVOT-FILTERS` for a hypothetical gradebook.
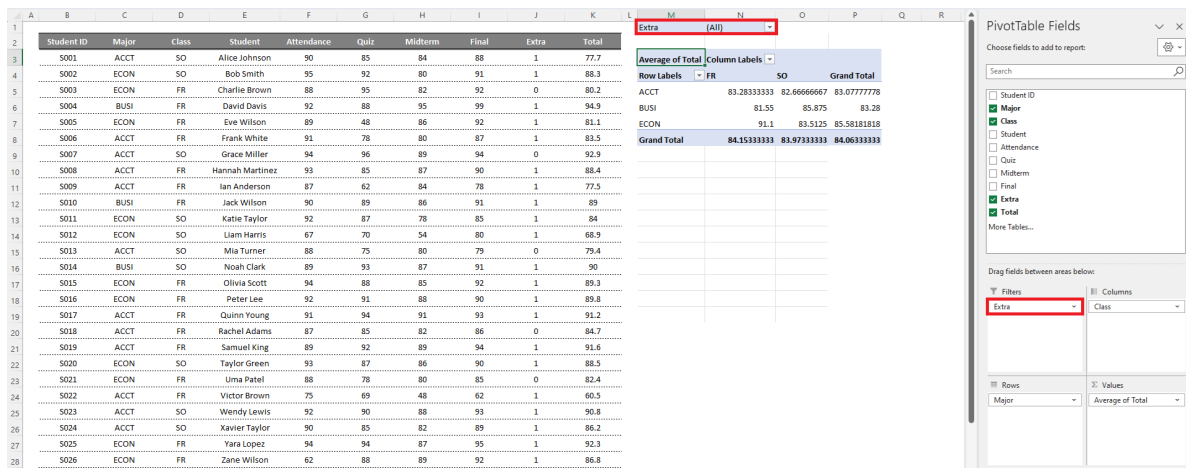

Figure 6: PivotTable Filters

This new filter in the red box allows us to filter the raw data so that only certain portions of the data is visible on the PivotTable. In this specific example, the filter is based on whether a student has extra credit or not.

If we select the value 0 in the filter shown in the blue box in Figure 7, the PivotTable will be show results based on only on the students without extra credit. If we select 1 in the filter, then the PivotTable will be calculated based on only the students who have extra credit.
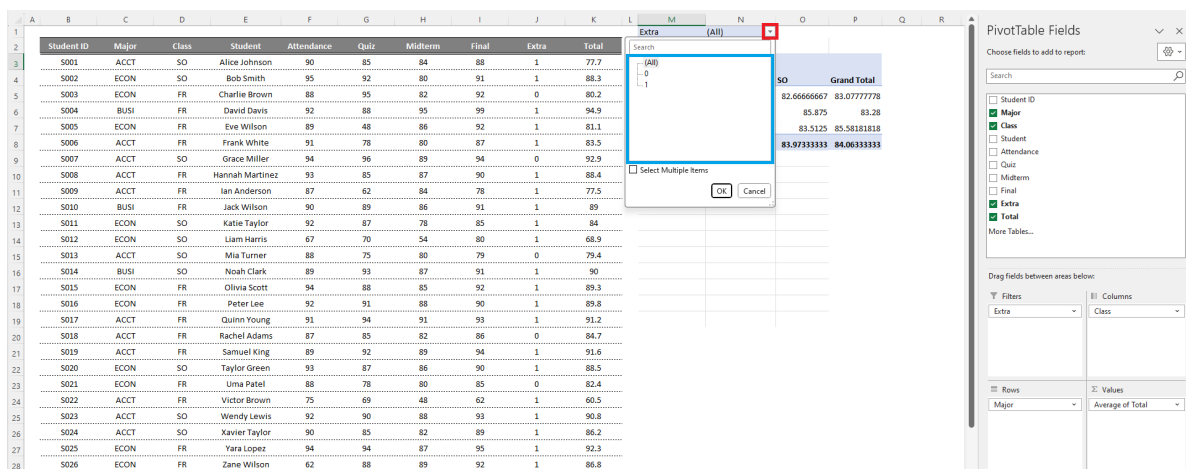
Figure 7: PivotTable Filtering

4

Other Filters on PivotTables

You may have also noticed that there are some filters that appear on the PivotTables by default. These can be seen in the red boxes in Figure 8. These filters can be used to remove certain variables, or sort the order of the variables on the PivotTable itself.
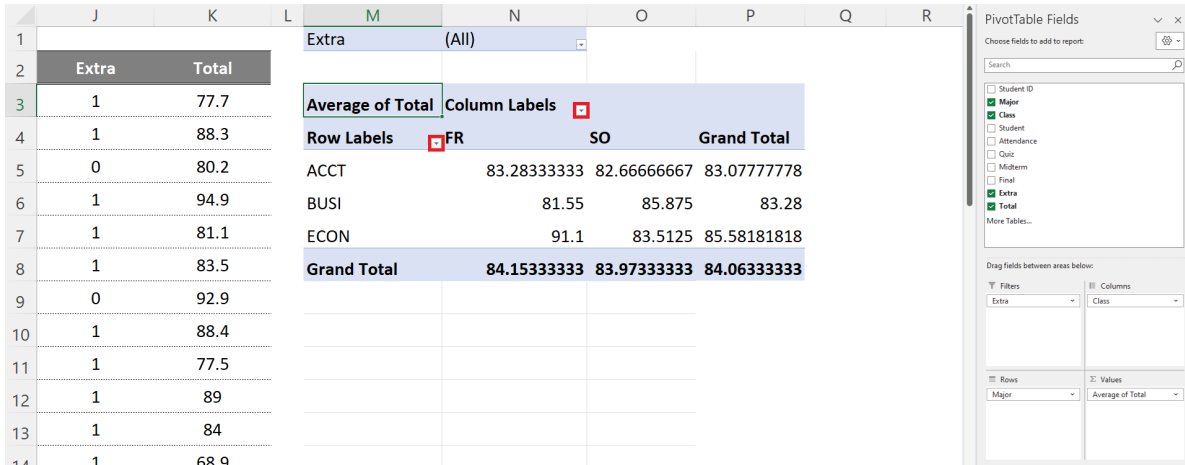

Figure 8: Other Filters in PivotTables

Topic 3. PivotTables: Calculated Fields

Sometimes, we may be interested in variables that are not already included in the source data. We can either edit the original data, but this editing the raw data manually is usually not considered "best practice," as there is an associated risk of contamination. Instead, if the data is easily calculated from the existing variables, we can use `Calculated Fields` to add a variable in the PivotTable.

Navigate to worksheet `PIVOT-CREATE`, which has synthetic data on the housing market near Chicago. Suppose you wanted to create a new variable that tells us the average price per square feet.
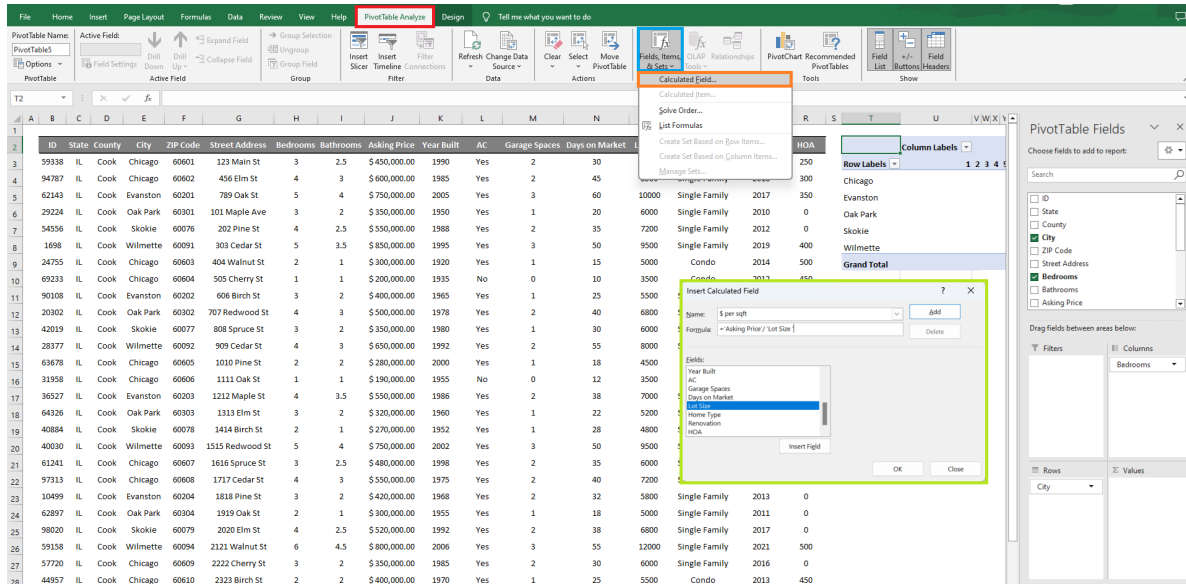
Figure 9: Calculated Fields

After creating an appropriate PivotTable, left click on any cell in the PivotTable, navigate to the `PivotTable Analysis` tab that is made available after creating and selecting a PivotTable. Then, select `Fields, Items, & Sets`, and then `Calculated Fields`. This will result in a new pop-up window in the green box where we can specify the new variable we want to add.

In this case, we wanted to add the asking price per square foot of the property. So, we first change the name to match the description of the variable. Then we can set up the formula either by typing it directly, or selecting them from the list of variables given in the pop-up window.

Figure 10: Calculated Fields with Errors

Resolving the #DIV/0! **Error**

In Figure 10, you will notice that there are a number of #DIV/0! errors. For instance, the value for one bedroom apartments in Evanston is #DIV/0!. This happened because there are no single bedroom apartments in Evanston, and thus Excel tried to fill the slot with $\frac{0}{0}$, which is undefined.

Figure 11: Correcting Errors in PivotTables

Left click on any of the #DIV/0! cells, right click, and then select PivotTable Options. In the new pop-up window, select the Layout & Format tab, and edit the options in the orange box. If you leave the box blank, any cells with error outputs will return an empty cell.

Figure 12: Corrected Errors in PivotTables

Topic 4. PivotTables: "Dividing" PivotTables

The use of filters in PivotTables also allows you to generate PivotTables across multiple sheets each corresponding to an item on the filter. Lets update the PivotTable we have in worksheet PIVOT-CREATE to the form you see in Figure 13. There were 5 cities in our data, and suppose we want to create a Pivot-Table telling us what the average price per sqft is for houses with 1,2,···,6 bedrooms in each city.

Figure 13: New PivotTable in PIVOT-CREATE

Navigate to the PivotTable Analyze, and click on Options, and then select Show Report Filter Pages. This will automatically create 5 worksheets with a mini PivotTable of each city in the dataset.

8