Lecture Note #17: PivotTables Part #3

BUSI 201: Business Data Analysis

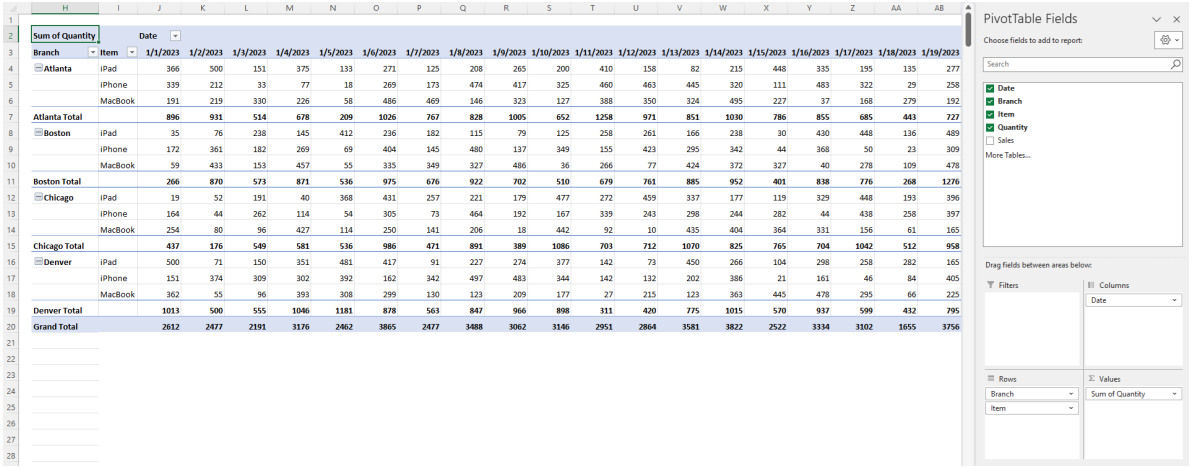
Fall 2023

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.



Sum of Quantity		Date																		
Branch	Item	1/1/2023	1/2/2023	1/3/2023	1/4/2023	1/5/2023	1/6/2023	1/7/2023	1/8/2023	1/9/2023	1/10/2023	1/11/2023	1/12/2023	1/13/2023	1/14/2023	1/15/2023	1/16/2023	1/17/2023	1/18/2023	1/19/2023
Atlanta	iPad	366	500	151	375	133	271	125	208	265	200	410	158	82	215	448	335	195	135	277
	iPhone	339	212	33	77	18	269	173	474	417	325	460	463	445	320	111	483	322	29	258
	MacBook	191	219	330	226	58	486	469	146	323	127	388	350	324	495	227	37	168	279	192
Atlanta Total		**896**	**931**	**514**	**678**	**209**	**1026**	**767**	**828**	**1005**	**652**	**1258**	**971**	**851**	**1030**	**786**	**855**	**685**	**443**	**727**
Boston	iPad	35	76	238	145	412	236	182	115	79	125	258	261	166	238	30	430	448	136	489
	iPhone	172	361	182	269	69	404	145	480	137	349	155	423	295	342	44	368	50	23	309
	MacBook	59	433	153	457	55	335	349	327	486	36	266	77	424	372	327	40	278	109	478
Boston Total		**266**	**870**	**573**	**871**	**536**	**975**	**676**	**922**	**702**	**510**	**679**	**761**	**885**	**952**	**401**	**838**	**776**	**268**	**1276**
Chicago	iPad	19	52	191	40	368	431	257	221	179	477	272	459	337	177	119	329	448	193	396
	iPhone	164	44	262	114	54	305	73	464	192	167	339	243	298	244	282	44	438	258	397
	MacBook	254	80	96	427	114	250	141	206	18	442	92	10	435	404	364	331	156	61	165
Chicago Total		**437**	**176**	**549**	**581**	**536**	**986**	**471**	**891**	**389**	**1086**	**703**	**712**	**1070**	**825**	**765**	**704**	**1042**	**512**	**958**
Denver	iPad	500	71	150	351	481	417	91	227	274	377	142	73	450	266	104	298	258	282	165
	iPhone	151	374	309	302	392	162	342	497	483	344	142	132	202	386	21	161	46	84	405
	MacBook	362	55	96	393	308	299	130	123	209	177	27	215	123	363	445	478	295	66	225
Denver Total		**1013**	**500**	**555**	**1046**	**1181**	**878**	**563**	**847**	**966**	**898**	**311**	**420**	**775**	**1015**	**570**	**937**	**599**	**432**	**795**
Grand Total		**2612**	**2477**	**2191**	**3176**	**2462**	**3865**	**2477**	**3488**	**3062**	**3146**	**2951**	**2864**	**3581**	**3822**	**2522**	**3334**	**3102**	**1655**	**3756**

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.

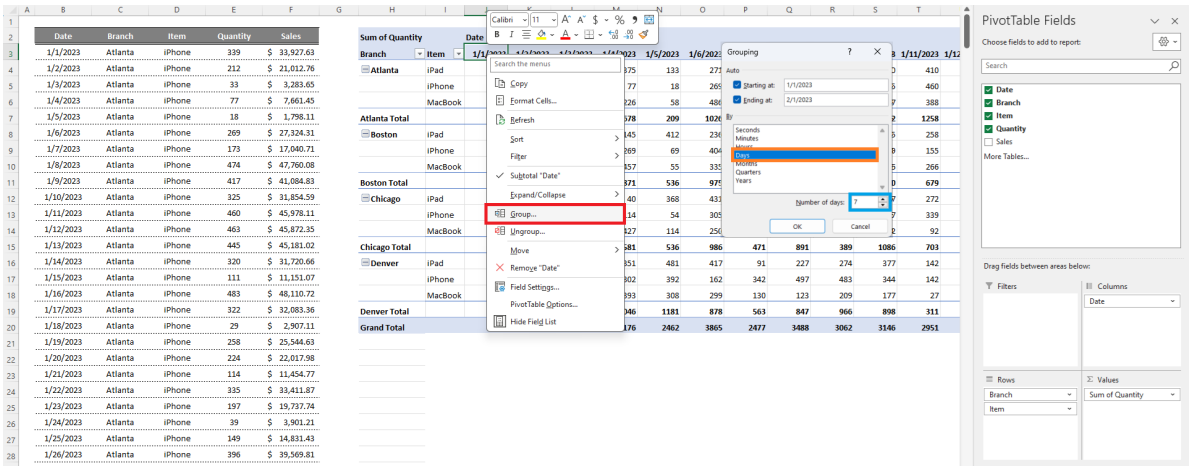


Figure 2: Grouping by Dates

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."

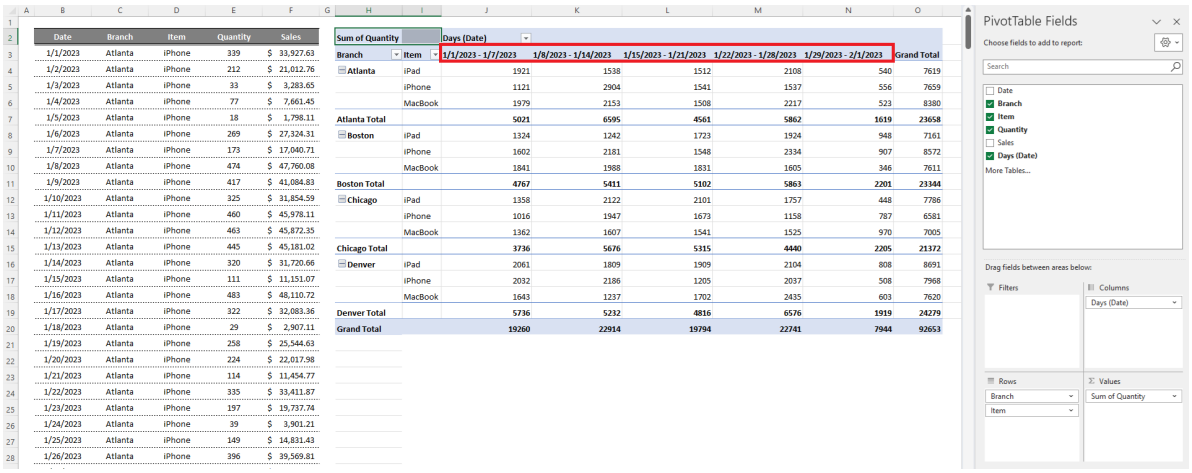


Figure 3: Grouped by Dates

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.

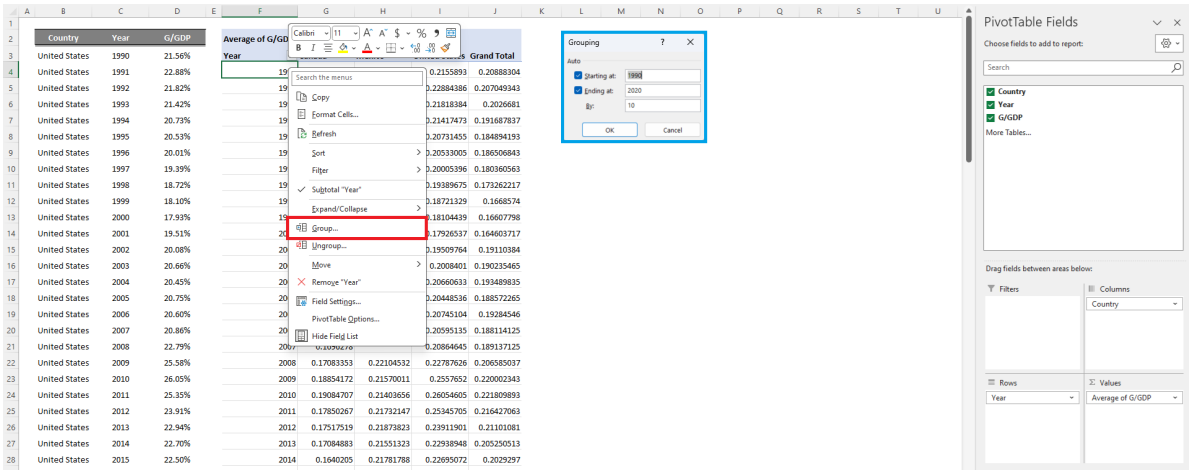


Figure 4: Grouping by Numbers

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.

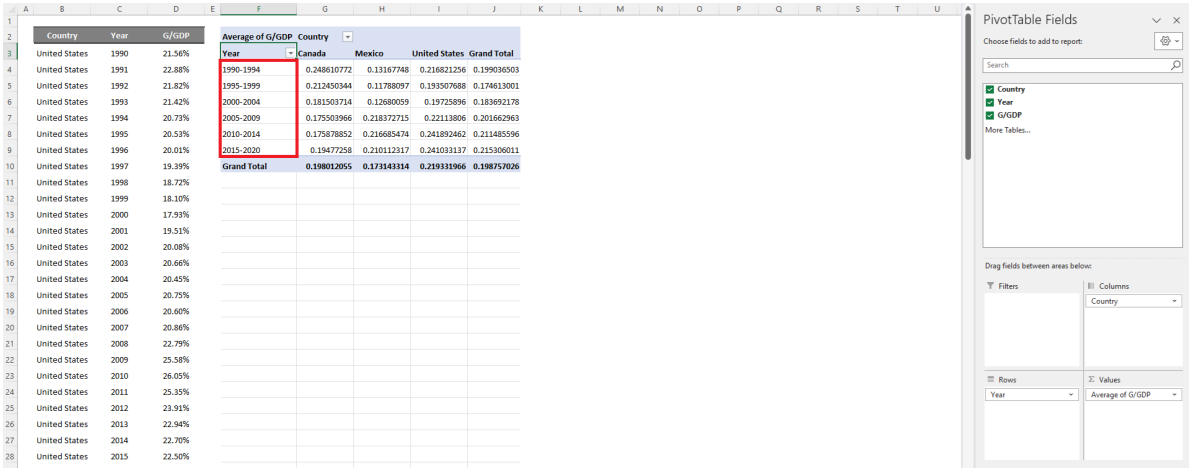
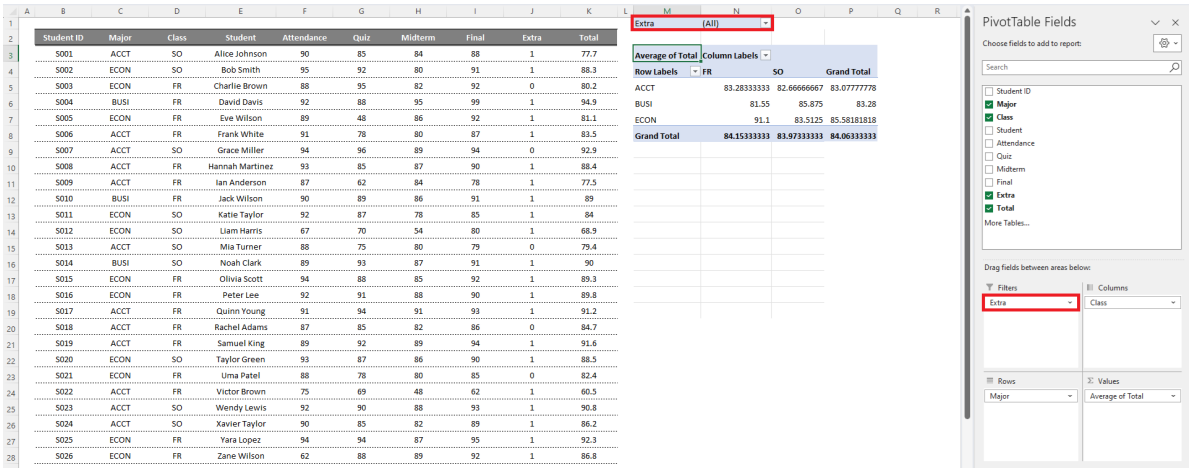


Figure 5: Grouped by Numbers

Topic 2. PivotTables: Filters

Up to this point, we have been primarily interested in three "boxes" when we were creating PivotTables; 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.

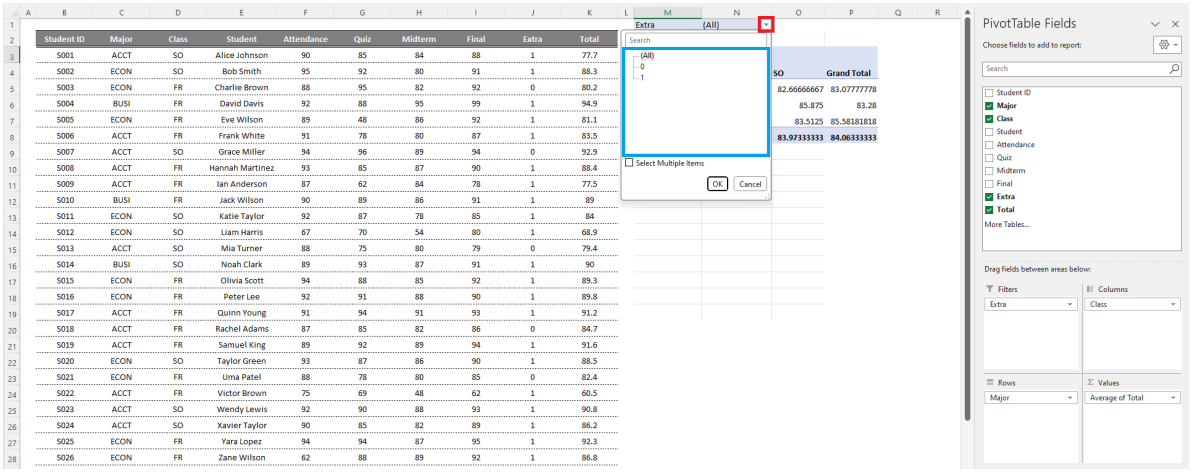


Student ID	Major	Class	Student	Attendance	Quiz	Midterm	Final	Extra	Total
S001	ACCT	SO	Alice Johnson	90	85	84	88	1	77.7
S002	ECON	SO	Bob Smith	95	92	80	91	1	88.3
S003	ECON	FR	Charlie Brown	88	95	82	92	0	80.2
S004	BUSI	FR	David Davis	92	88	95	99	1	94.9
S005	ECON	FR	Eve Wilson	89	48	86	92	1	81.1
S006	ACCT	FR	Frank White	91	78	80	87	1	83.5
S007	ACCT	SO	Grace Miller	94	96	89	94	0	92.9
S008	ACCT	FR	Hannah Martinez	93	85	87	90	1	88.4
S009	ACCT	FR	Ian Anderson	87	62	84	78	1	77.5
S010	BUSI	FR	Jack Wilson	90	89	86	91	1	89
S011	ECON	SO	Katie Taylor	92	87	78	85	1	84
S012	ECON	SO	Liam Harris	67	70	54	80	1	68.9
S013	ACCT	SO	Mia Turner	88	75	80	79	0	79.4
S014	BUSI	SO	Noah Clark	89	93	87	91	1	90
S015	ECON	FR	Olivia Scott	94	88	85	92	1	89.3
S016	ECON	FR	Peter Lee	92	91	88	90	1	89.8
S017	ACCT	FR	Quinn Young	91	94	91	93	1	91.2
S018	ACCT	FR	Rachel Adams	87	85	82	86	0	84.7
S019	ACCT	FR	Samuel King	89	92	89	94	1	91.6
S020	ECON	SO	Taylor Green	93	87	86	90	1	88.5
S021	ECON	FR	Uma Patel	88	78	80	85	0	82.4
S022	ACCT	FR	Victor Brown	75	69	48	62	1	60.5
S023	ACCT	SO	Wendy Lewis	92	90	88	93	1	90.8
S024	ACCT	SO	Xavier Taylor	90	85	82	89	1	86.2
S025	ECON	FR	Yara Lopez	94	94	87	95	1	92.3
S026	ECON	FR	Zane Wilson	62	88	89	92	1	86.8

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.

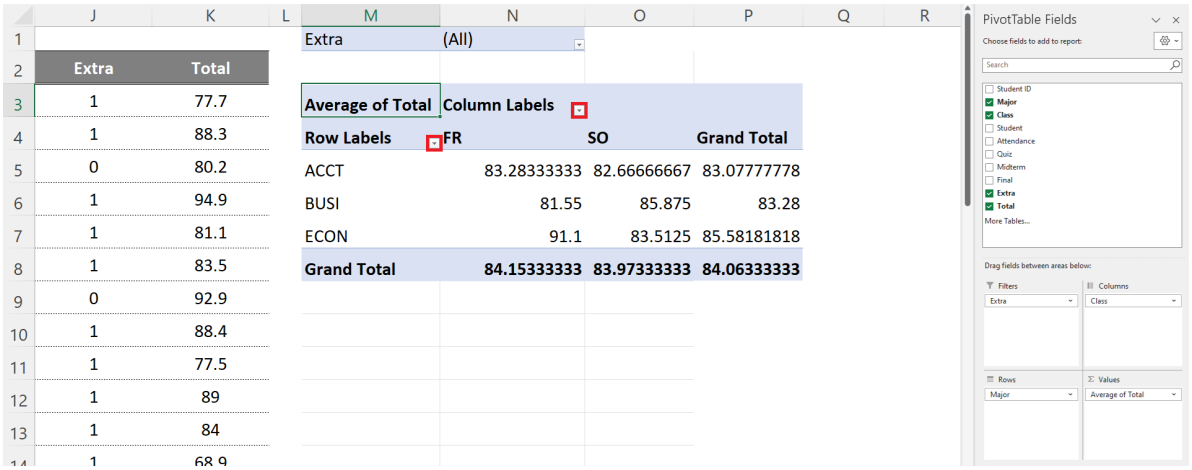


Student ID	Major	Class	Student	Attendance	Quiz	Midterm	Final	Extra	Total
S001	ACCT	SO	Alice Johnson	90	85	84	88	1	77.7
S002	ECON	SO	Bob Smith	95	92	80	91	1	88.3
S003	ECON	FR	Charlie Brown	88	95	82	92	0	80.2
S004	BUSI	FR	David Davis	92	88	95	99	1	94.9
S005	ECON	FR	Eve Wilson	89	48	86	92	1	81.1
S006	ACCT	FR	Frank White	91	78	80	87	1	83.5
S007	ACCT	SO	Grace Miller	94	96	89	94	0	92.9
S008	ACCT	FR	Hannah Martinez	93	85	87	90	1	88.4
S009	ACCT	FR	Ian Anderson	87	62	84	78	1	77.5
S010	BUSI	FR	Jack Wilson	90	89	86	91	1	89
S011	ECON	SO	Katie Taylor	92	87	78	85	1	84
S012	ECON	SO	Liam Harris	67	70	54	80	1	68.9
S013	ACCT	SO	Mia Turner	88	75	80	79	0	79.4
S014	BUSI	SO	Noah Clark	89	93	87	91	1	90
S015	ECON	FR	Olivia Scott	94	88	85	92	1	89.3
S016	ECON	FR	Peter Lee	92	91	88	90	1	89.8
S017	ACCT	FR	Quinn Young	91	94	91	93	1	91.2
S018	ACCT	FR	Rachel Adams	87	85	82	86	0	84.7
S019	ACCT	FR	Samuel King	89	92	89	94	1	91.6
S020	ECON	SO	Taylor Green	93	87	86	90	1	88.5
S021	ECON	FR	Uma Patel	88	78	80	85	0	82.4
S022	ACCT	FR	Victor Brown	75	69	48	62	1	60.5
S023	ACCT	SO	Wendy Lewis	92	90	88	93	1	90.8
S024	ACCT	SO	Xavier Taylor	90	85	82	89	1	86.2
S025	ECON	FR	Yara Lopez	94	94	87	95	1	92.3
S026	ECON	FR	Zane Wilson	62	88	89	92	1	86.8

Figure 7: PivotTable Filtering

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.



Extra	Total
1	77.7
1	88.3
0	80.2
1	94.9
1	81.1
1	83.5
0	92.9
1	88.4
1	77.5
1	89
1	84
1	68.9

Extra (All)

Average of Total	Column Labels		
Row Labels	FR	SO	Grand Total
ACCT	83.28333333	82.66666667	83.07777778
BUSI	81.55	85.875	83.28
ECON	91.1	83.5125	85.58181818
Grand Total	**84.15333333**	**83.97333333**	**84.06333333**

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.

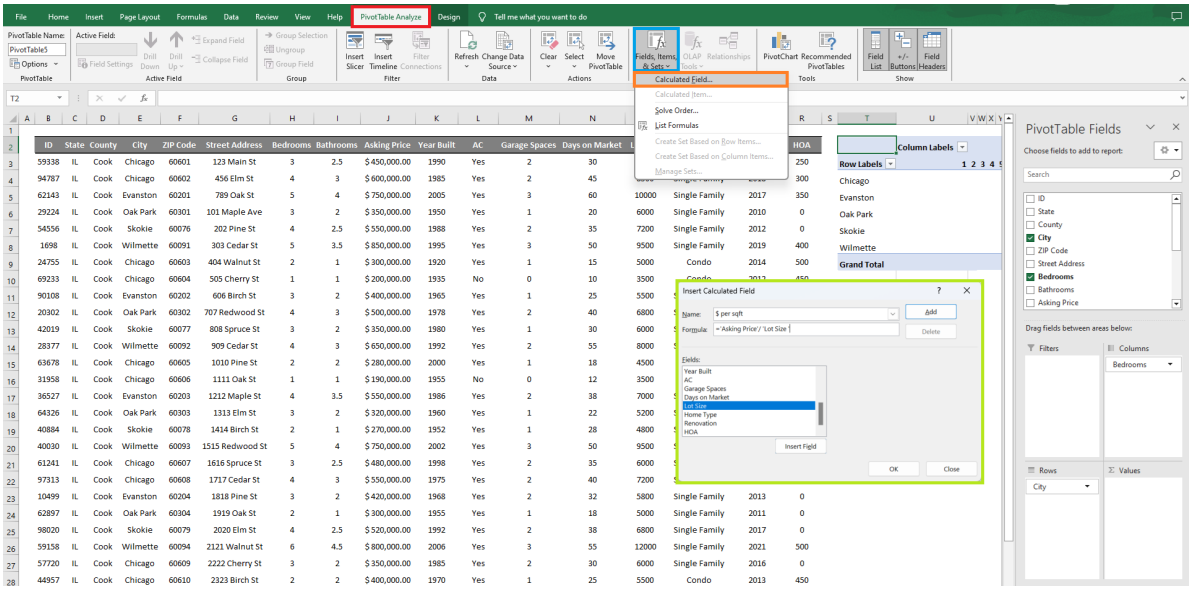
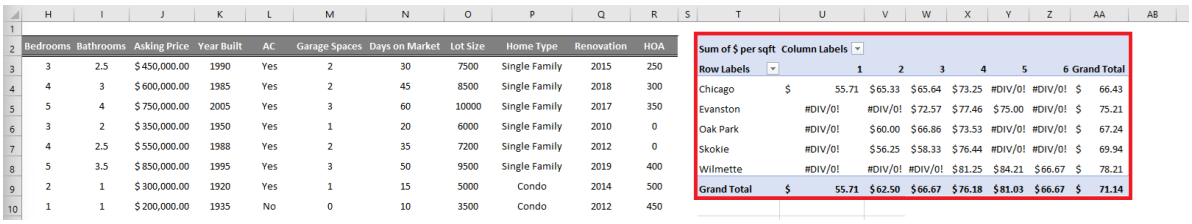


Figure 9: Calculated Fields

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.



Bedrooms	Bathrooms	Asking Price	Year Built	AC	Garage Spaces	Days on Market	Lot Size	Home Type	Renovation	HOA
3	2.5	$450,000.00	1990	Yes	2	30	7500	Single Family	2015	250
4	3	$600,000.00	1985	Yes	2	45	8500	Single Family	2018	300
5	4	$750,000.00	2005	Yes	3	60	10000	Single Family	2017	350
3	2	$350,000.00	1950	Yes	1	20	6000	Single Family	2010	0
4	2.5	$550,000.00	1988	Yes	2	35	7200	Single Family	2012	0
5	3.5	$850,000.00	1995	Yes	3	50	9500	Single Family	2019	400
2	1	$300,000.00	1920	Yes	1	15	5000	Condo	2014	500
1	1	$200,000.00	1935	No	0	10	3500	Condo	2012	450

Sum of $ per sqft	Column Labels						
Row Labels	1	2	3	4	5	6	Grand Total
Chicago	$ 55.71	$65.33	$65.64	$73.25	#DIV/0!	#DIV/0!	$ 66.43
Evanston	#DIV/0!	#DIV/0!	$72.57	$77.46	$75.00	#DIV/0!	$ 75.21
Oak Park	#DIV/0!	$60.00	$66.86	$73.53	#DIV/0!	#DIV/0!	$ 67.24
Skokie	#DIV/0!	$56.25	$58.33	$76.44	#DIV/0!	#DIV/0!	$ 69.94
Wilmette	#DIV/0!	#DIV/0!	#DIV/0!	$81.25	$84.21	$66.67	$ 78.21
Grand Total	$ 55.71	$62.50	$66.67	$76.18	$81.03	$66.67	$ 71.14

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.

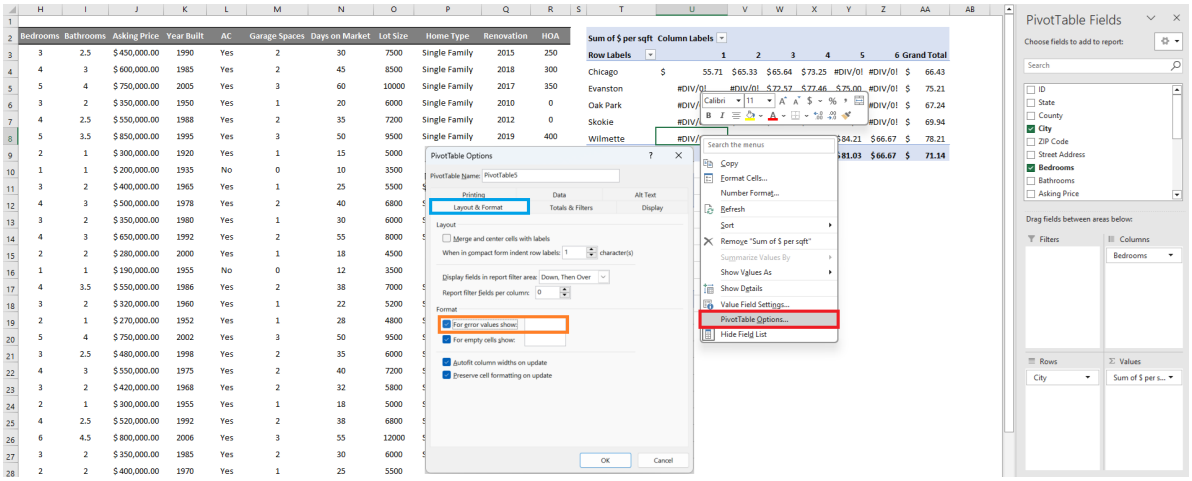


Figure 11: Correcting Errors in PivotTables

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.

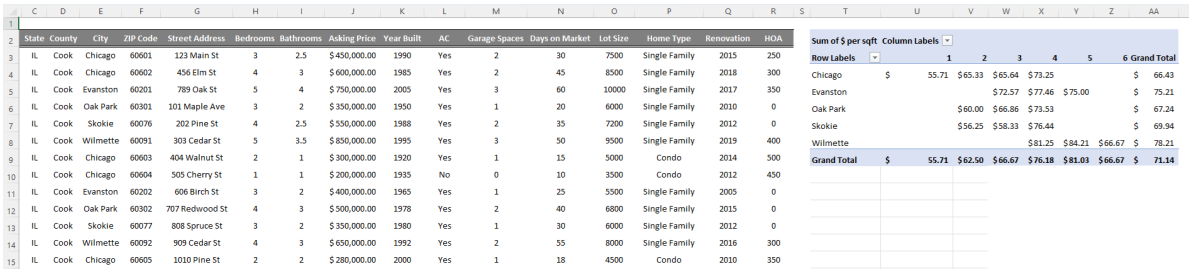


Figure 12: Corrected Errors in PivotTables

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 PivotTable telling us what the average price per sqft is for houses with 1,2,⋯,6 bedrooms in each city.

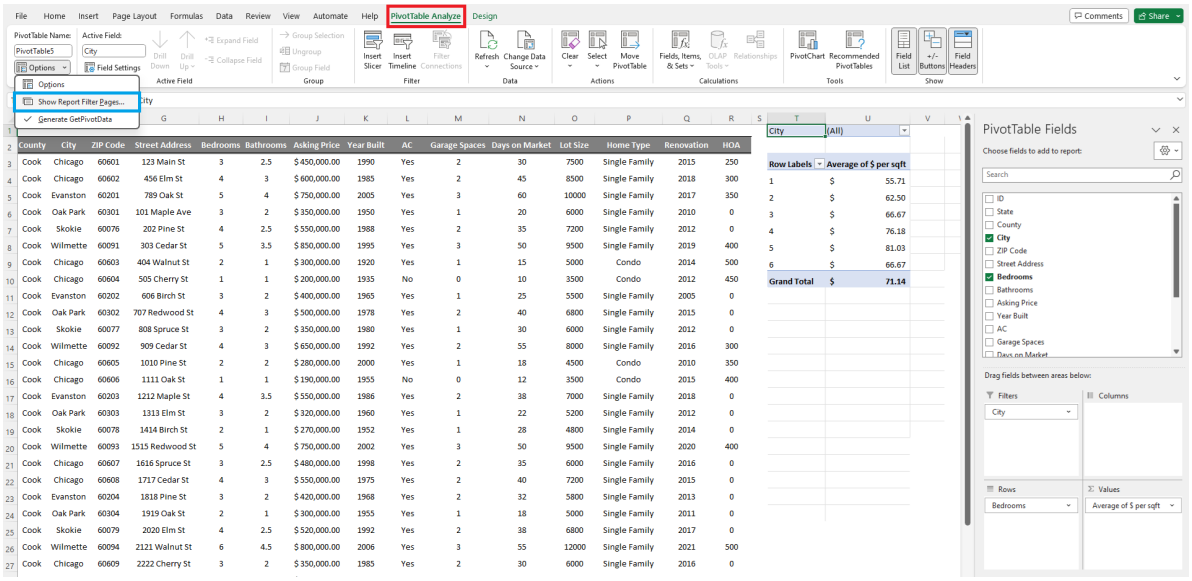


Figure 13: New PivotTable in `PIVOT-CREATE`

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.