

Intro to Data Science

Basic Business Analytics

1	KEY METRICS ANALYSIS	1
1.1	PRACTICAL EXERCISE FOR KEY METRICS ANALYSIS.....	8
2	COMPARISON ANALYSIS.....	9
2.1	PRACTICAL EXERCISE FOR COMPARISON ANALYSIS.....	14
3	TREND ANALYSIS	15
3.1	PRACTICAL EXERCISE FOR TREND ANALYSIS.....	21
4	RANKING ANALYSIS.....	22
4.1	PRACTICAL EXERCISE FOR RANKING ANALYSIS	27

1 Key Metrics Analysis

Key Metrics or Key Performance Indicators (KPIs) are the main method that we use to measure the performance of a company, division, department or team. Key Metrics provide an overall understanding of how the business is performing. Key Metrics usually include the following types of questions:

- a) What is the Total Sales?
- b) What is the Total Profit?
- c) What is the Profit Ratio?
- d) How many transactions were there processed?
- e) What was the average Sales amount?
- f) What was the highest Sales amount?
- g) What was the lowest Sales amount?

File to use: SalesData-v1.xlsx

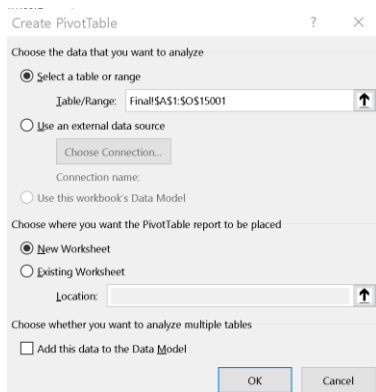
Create a Pivot table from the current range of data.

Make sure you have selected a cell in the range.

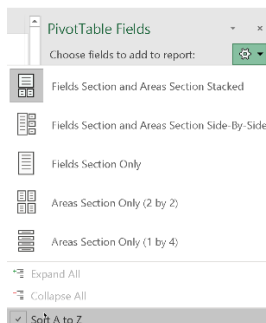
Insert -> Pivot table

The Create Pivot table should automatically detect the complete range of cells with valid values.

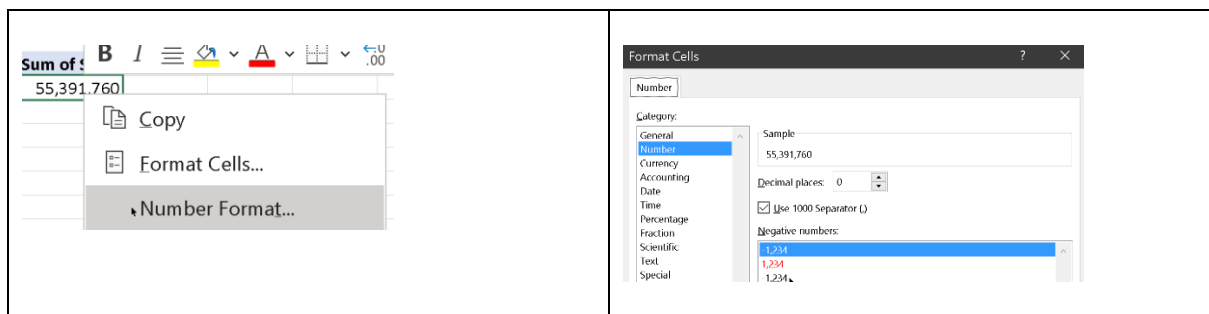
Make sure you create the Pivot table in a new worksheet.



You can change the order of the Pivot Table fields as they appear in the main list: its sometimes useful to sort them in alphabetical order if you have many fields.



Select the Sales field to get the Sum of Sales, then perform a Number Format to make it easier to view.



Repeat this with the Profit field to get the Sum of Sales, with a similar Number Format as well.

Sum of Sales	Sum of Profit
55,391,760	35,589,375

Next insert a Calculated Field to obtain the Profit Ratio, and format the result accordingly

--	--

The screenshot shows two Excel dialog boxes. On the left, the 'Insert Calculated Field' dialog has 'Name: Profit Ratio' and 'Formula: = Profit / Sales'. The 'Fields' list on the left includes 'Order Date', 'Year', 'Order Qty', 'Cost of Sales', 'Sales', 'Profit', 'Channel', and 'Product Name'. On the right, the 'Format Cells' dialog shows the 'Number' category with 'Percentage' selected, displaying a sample of '64.3%' and 'Decimal places: 1'.

Sum of Sales	Sum of Profit	Sum of Profit Ratio
55,391,760	35,589,375	64.3%

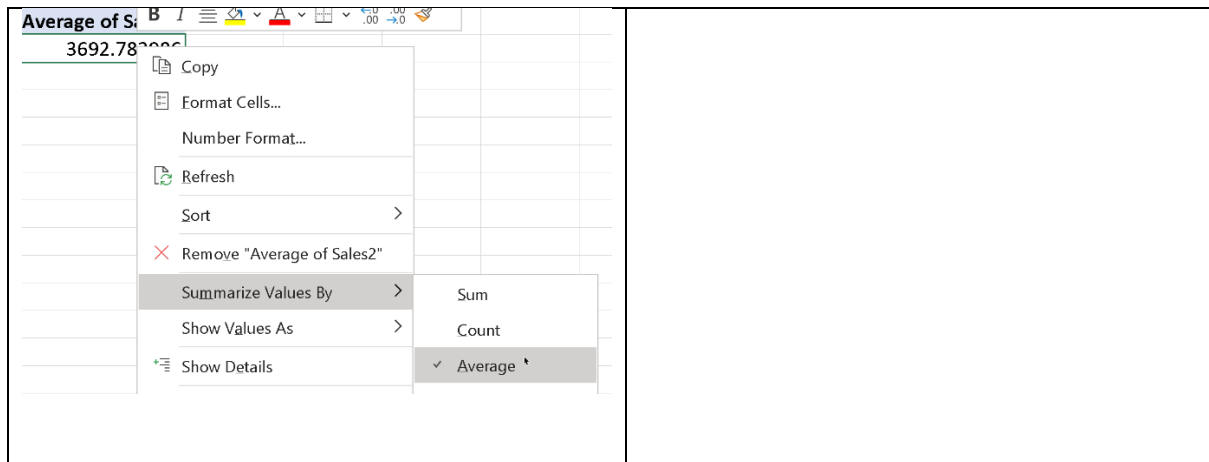
To count the number of transactions, we would typically use a unique Sales ID, which we don't have here in this dataset. We can initially use the Sales Field to start with (Drag and drop into the Values area), but this time change it to Summarize Values by Count to get the actual number of transactions (which we can also double confirm from the source dataset)

The screenshot shows two parts. On the left, a context menu for 'Count of Sales' is open, showing options like 'Copy', 'Format Cells...', 'Refresh', 'Sort', 'Remove "Count of Sales2"', 'Summarize Values By' (set to 'Sum'), and 'Show Values As' (set to 'Count'). On the right, a table shows the results of summarizing by count:

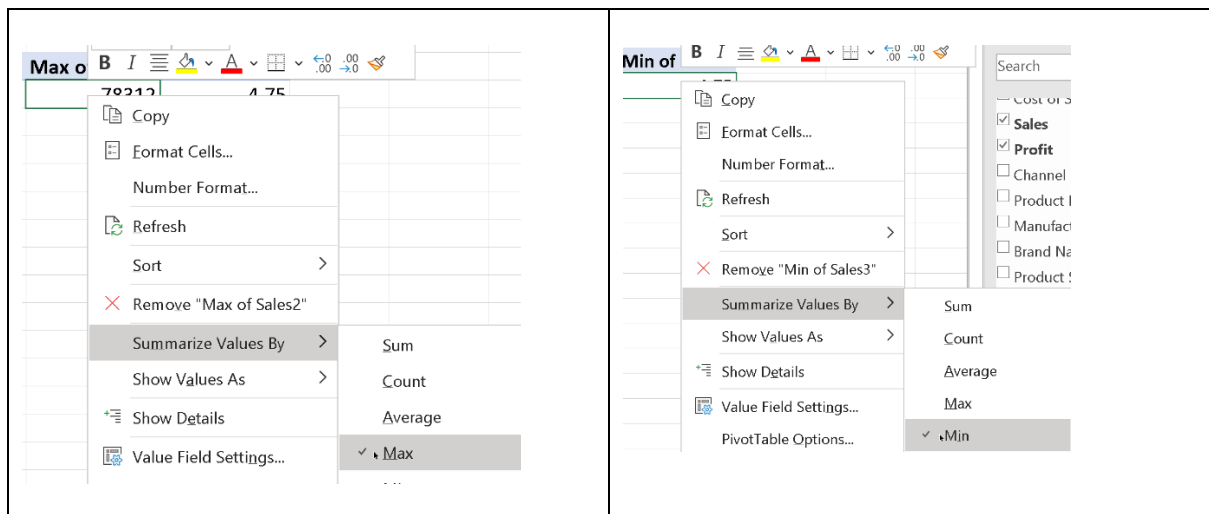
Sum of Sales	Sum of Profit	Sum of Profit Ratio	Count of Sales2
55,391,760	35,589,375	64.3%	15000

To get the average Sales, use the Sales field again, but this time Use Summarize Values by Average, and format it accordingly.

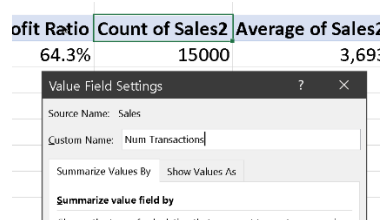
Count of Sales2	Average of Sales2
15000	3,693



Add two more Sales Fields into the Values area, and then Summarize Values by Min and Max for these two fields and format them appropriately.



Now that we have all our numbers, you can double click on any (or all) of the field headers to change the field names to something more appropriate if you wish.



We now have our key metrics listed out in a row:

Sum of Sales	Sum of Profit	Sum of Profit Ratio	Num Transactions	Average of Sales	Highest Sales	Lowest Sales
55,391,760	35,589,375	64.3%	15000	3,693	78312	5

You can transport the rows and columns for a better layout:

--	--

<p>Drag fields between areas below:</p> <div> <div>Filters</div> <div>Columns</div> </div> <div> <div>Rows</div> <div>Σ Values</div> </div> <div> <div>Σ Values</div> <div>Sum of Sales</div> <div>Sum of Profit</div> <div>Sum of Profit Ratio</div> </div>	<table> <thead> <tr> <th colspan="2">Values</th></tr> </thead> <tbody> <tr> <td>Total Sales</td><td>55,391,760</td></tr> <tr> <td>Total Profit</td><td>35,589,375</td></tr> <tr> <td>Actual Profit Ratio</td><td>64.3%</td></tr> <tr> <td>Num Transactions</td><td>15000</td></tr> <tr> <td>Average of Sales</td><td>3,693</td></tr> <tr> <td>Highest Sales</td><td>78312</td></tr> <tr> <td>Lowest Sales</td><td>5</td></tr> </tbody> </table>	Values		Total Sales	55,391,760	Total Profit	35,589,375	Actual Profit Ratio	64.3%	Num Transactions	15000	Average of Sales	3,693	Highest Sales	78312	Lowest Sales	5
Values																	
Total Sales	55,391,760																
Total Profit	35,589,375																
Actual Profit Ratio	64.3%																
Num Transactions	15000																
Average of Sales	3,693																
Highest Sales	78312																
Lowest Sales	5																

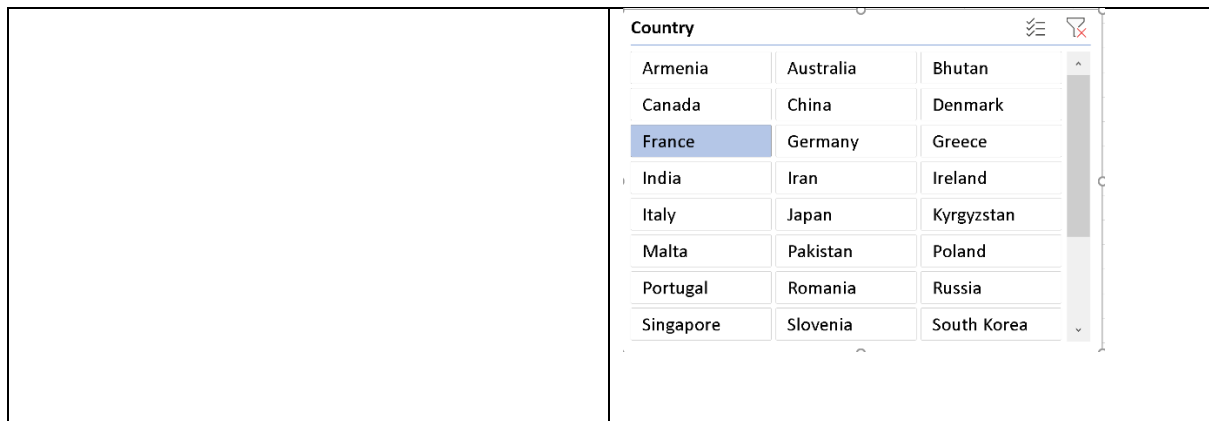
To zoom in and get more details on the metrics based on a certain category (for e.g. country), we can filter the table to decide which subset of records to apply the various aggregation operations we have performed so far on.

For this purpose, we can use a slicer. Go to Pivot Table Analyze -> Insert Slicer, and select Country. You can then select particular countries from the slicer drop down to get the metric values only for sales records related to that particular country.

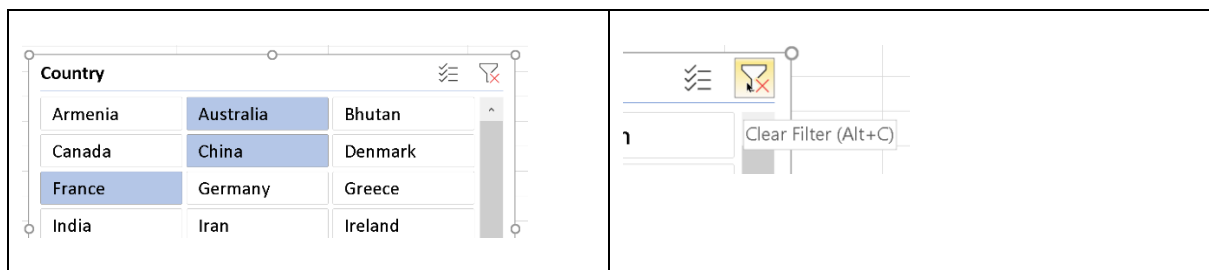
<p>Review View Help Analyze Design</p> <p>Insert Slicer Timeline Connections Refresh</p> <p>Insert Slicers</p> <ul style="list-style-type: none"> <input type="checkbox"/> Order Date <input type="checkbox"/> Year <input type="checkbox"/> Order Qty <input type="checkbox"/> Cost of Sales <input type="checkbox"/> Sales <input type="checkbox"/> Profit <input type="checkbox"/> Channel <input type="checkbox"/> Product Name <input type="checkbox"/> Manufacturer <input type="checkbox"/> Brand Name <input type="checkbox"/> Product Sub Category <input type="checkbox"/> Product Category <input type="checkbox"/> Region <input type="checkbox"/> City <input checked="" type="checkbox"/> Country 	<table> <thead> <tr> <th colspan="2">Values</th></tr> </thead> <tbody> <tr> <td>Total Sales</td><td>1,148,316</td></tr> <tr> <td>Total Profit</td><td>694,190</td></tr> <tr> <td>Actual Profit Ratio</td><td>60.5%</td></tr> <tr> <td>Num Transactions</td><td>359</td></tr> <tr> <td>Average of Sales</td><td>3,199</td></tr> <tr> <td>Highest Sales</td><td>45900</td></tr> <tr> <td>Lowest Sales</td><td>37</td></tr> </tbody> </table> <p>Country</p> <ul style="list-style-type: none"> Armenia Australia Bhutan Canada China Denmark France Germany 	Values		Total Sales	1,148,316	Total Profit	694,190	Actual Profit Ratio	60.5%	Num Transactions	359	Average of Sales	3,199	Highest Sales	45900	Lowest Sales	37
Values																	
Total Sales	1,148,316																
Total Profit	694,190																
Actual Profit Ratio	60.5%																
Num Transactions	359																
Average of Sales	3,199																
Highest Sales	45900																
Lowest Sales	37																

You can use the Slicer Tools options to adjust the number of columns, as well as their height and width to make them easier to work with:

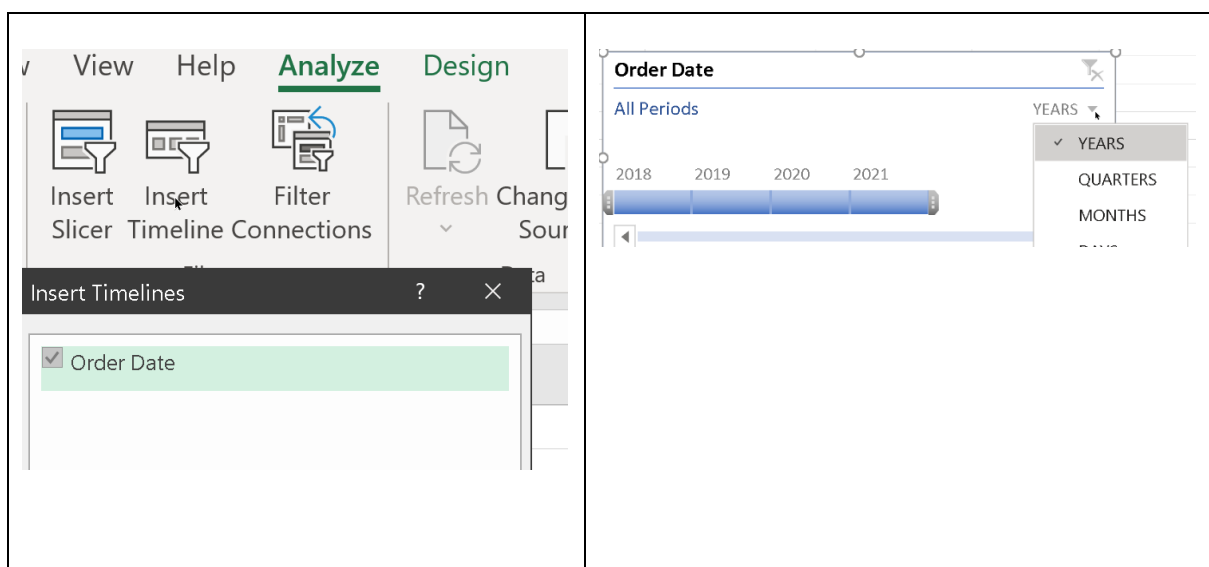
<p>Slicer Tools SalesData v1.xlsx</p> <p>Options</p> <p>Columns: 3</p> <p>Height: 7.01 cm</p> <p>Width: 9.87 cm</p>	
---	--



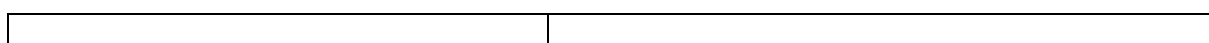
You can use the Ctrl-Key to select multiple countries at the same time, so the aggregation operations will apply to all these countries. You can then select the clear Filter button to clear the filter and make the aggregation apply to the entire dataset.

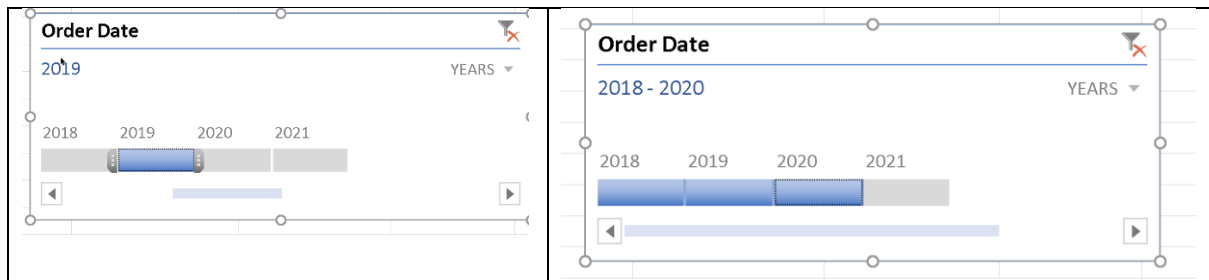


You can now also insert a Timeline and decide which particular period (Years, Quarters, Months) to examine.

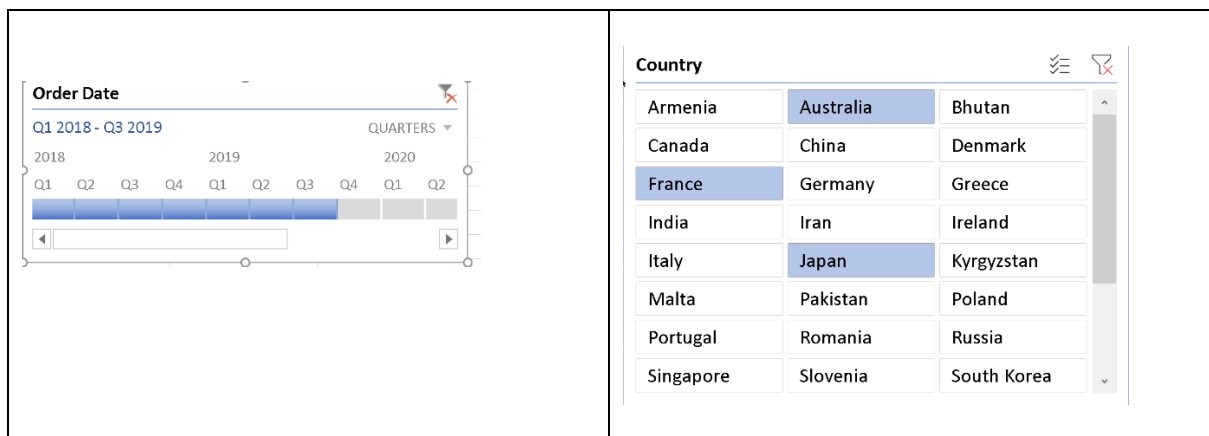


You can select a particular period or a range of periods (Shift-Click):

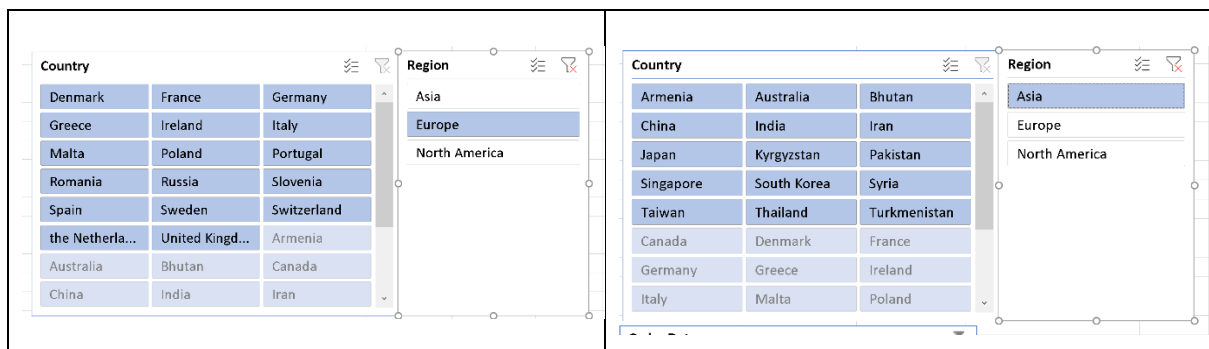




You can also select items from both filters to combine their filtering action together, for e.g. to perform aggregation operations on a specific country (or group of countries) over a specific period of time.

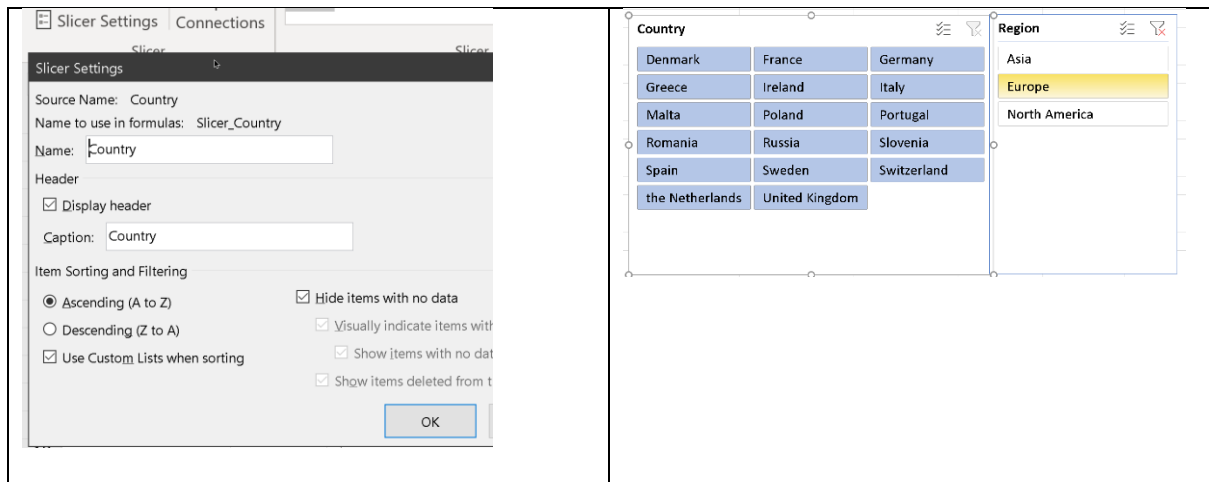


You can introduce another slicer for Region, and notice that selection of items in either slicer for Country or Region will influence the other: based on which countries are included in which region. To see this properly in effect, you will have to clear the filter in either one or both of the slicers.



You can also set the Slicer Settings to only clearly show the countries within a particular region, and not just grey out the countries that are not in that region (to make the analysis even more clearer):





1.1 Practical Exercise for Key Metrics Analysis

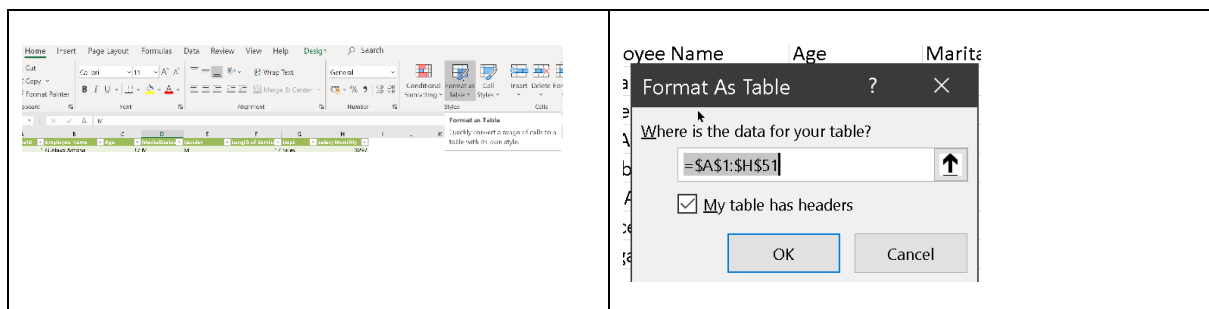
File to use: Employee-v1.xlsx

Analytical activities to perform:

1. How many employees are in John's company?
2. What is the average age of the employees?
3. What is the highest age at the company?
4. What is the lowest age at the company?
5. What is the average Length of Service?
6. What is the longest Length of Service?

Drill down further into this information based on the following categories: Dept and Gender

You can initially format the range of cells of the original data set as a table to make it easier to work with, and to create a Pivot Table. To do this, select a cell in the data range, go to Home -> Format as Table, format it and give it an appropriate name.



EmployeeID	EmployeeName	Age	MaritalStatus	Gender	Length of Service	Dept	Salary Monthly
1	Gustavo Ashong	37 M	M	F	17	Sales	3297
2	Catherine Abel	32 S	F	F	9	Sales	1859
3	Kim Abercrombie	54 M	F	F	11	Finance	4605
4	Humberto Acevedo	54 S	M	M	12	Logistics	1501
5	Pilar Ackerman	70 M	F	F	12	Human Resou	2916
6	Frances Adams	54 S	F	F	14	Sales	1820
7	Margaret Smith	73 S	F	F	6	Sales	3466
8	Carla Adams	73 M	F	F	7	Finance	3446
9	Jay Adams	26 M	M	M	13	Finance	4275
10	Ronald Adina	53 S	M	M	9	Logistics	4883

Once done, you can generate a Pivot Table in the usual way, but this time referencing the Table.

Create PivotTable

Choose the data that you want to analyze

☒ Select a table or range

Table/Range:

☐ Use an external data source

Connection name:

☐ Use this workbook's Data Model

Choose where you want the PivotTable report to be placed

☒ New Worksheet

☐ Existing Worksheet

Location:

Choose whether you want to analyze multiple tables

☐ Add this data to the Data Model

2 Comparison Analysis

Comparison analysis is probably one of the most popular forms of analysis you're going to do. The focus is finding the magnitude of difference between values for different fields. Visualizing your data set makes it much easier to appreciate this as opposed to directly examining the values themselves.

Sample analysis

- compare the sales and profit by year
- compare the sales by manufacturer
- compare the profit by product categories
- compare the sales by product category and by channel
- compare the profit by product subcategory and review that by different countries.
- compare the sales and profit by the countries, and review that by different product sub categories

File to use: SalesData-v2.xlsx

First, get the Sum of Sales and Profit by Years (Rows) and format the cells with an appropriate numeric format.

--	--

Drag fields between areas below:

Filters

Rows

Year

Columns

Σ Values

Σ Values

Sum of Sales

Sum of Profit

Row Labels	Sum of Sales	Sum of Profit
2018	18,919,151	12,057,185
2019	17,741,637	11,378,957
2020	17,284,124	11,237,878
2021	1,446,849	915,355
Grand Total	55,391,760	35,589,375

We can also sort on the Sum of Sales or Profit (either Smallest to Largest or vice versa)

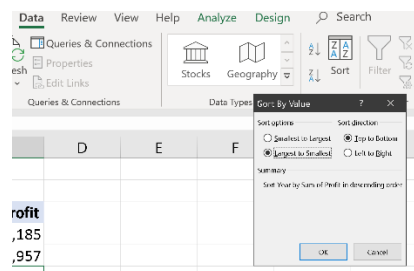
Row Labels	Sum of Sales	Sum of Profit
2018	18,919,151	12,057,185
2019	17,741,637	11,378,957
2020	17,284,124	11,237,878
2021	1,446,849	915,355
Grand Total	55,391,760	35,589,375

Sort

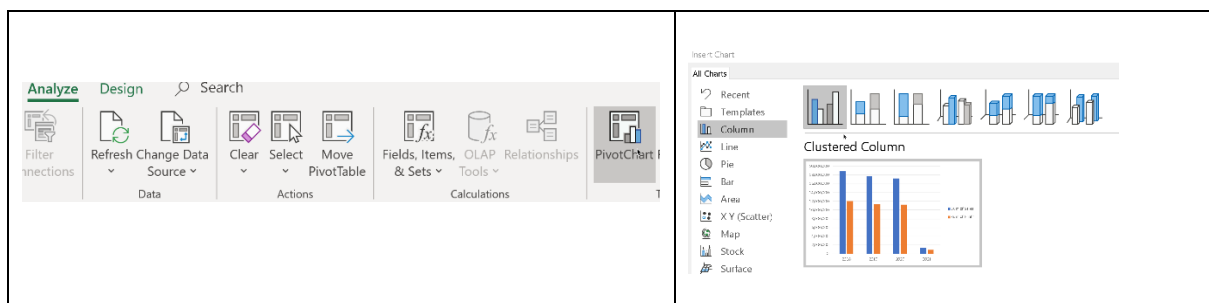
Sort Smallest to Largest

Sort Largest to Smallest

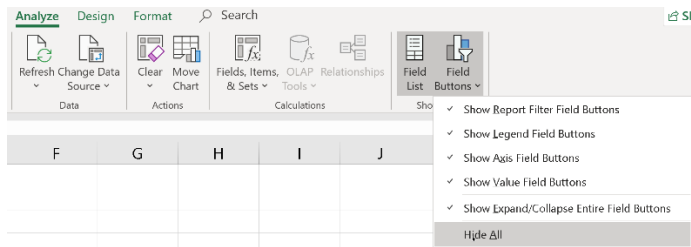
The sorting can also be done via the Sort option in the Data tab



Sort in descending order for Sum of Sales, and visualize this via a Pivot Chart (choose Clustered Column)



You can format what fields to be shown in the chart, including hiding everything:



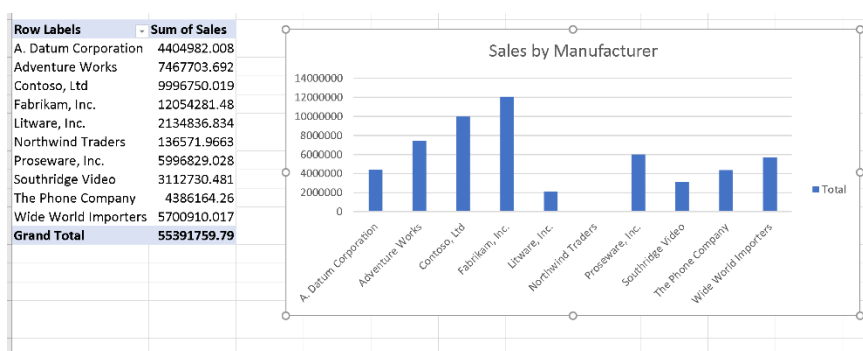
You can also change the Chart Elements appropriately, for e.g. add in Title and moving the legend to the Top.

We can now (using either the current pivot table or creating a new pivot table) repeat this earlier process to get the Sum of Sales by Manufacturer (Rows) and format the cells with an appropriate numeric format.

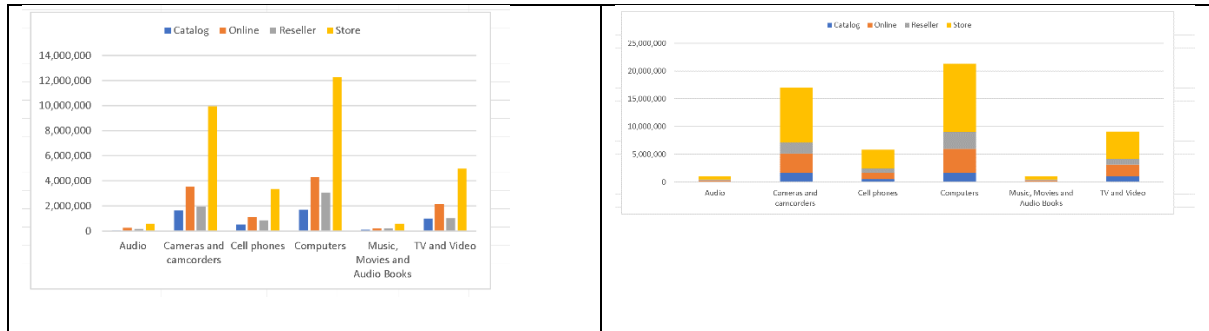
Rows	Values
Manufacturer	Sum of Sales

Row Labels	Sum of Sales
A. Datum Corporation	4404982.008
Adventure Works	7467703.692
Contoso, Ltd	9996750.019
Fabrikam, Inc.	12054281.48
Litware, Inc.	2134836.834
Northwind Traders	136571.9663
Proseware, Inc.	5996829.028
Southridge Video	3112730.481
The Phone Company	4386164.26
Wide World Importers	5700910.017
Grand Total	55391759.79

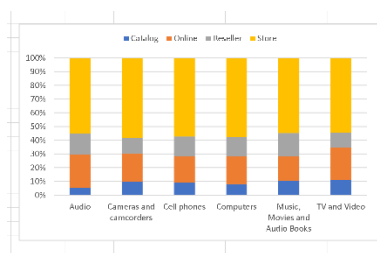
We can now insert another Clustered Column chart, and format it in the same way as we did previously.



We can also opt to change the Chart Type to another type for example. Clustered Bar.



We could also use the a 100% stacked column as well, which is useful for seeing how each of the different channels contribute for all products, regardless of the product total - this is useful for products whose sales totals are very small compared to other products and therefore will not come out clearly in the previous graphs.



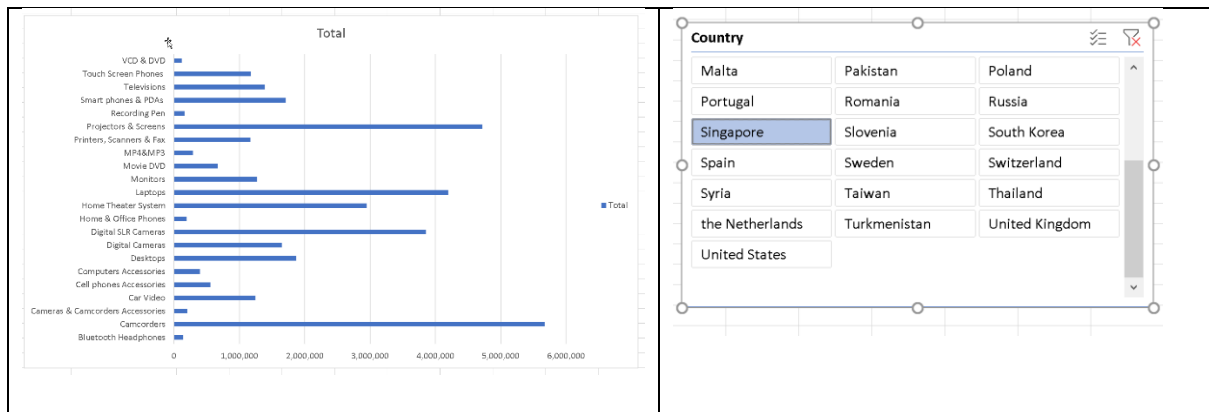
For the next analysis, we can (using either the current pivot table or creating a new pivot table) repeat this earlier process to get the Sum of Profit by Product Sub Category (Rows) and format the cells with an appropriate numeric format.

Axis (Categories)	Σ Values
Product Sub Category	Sum of Profit

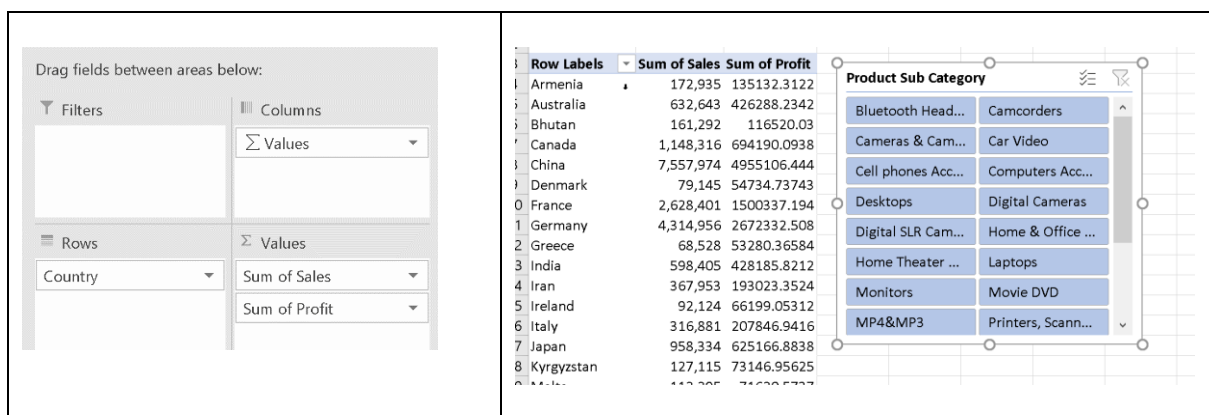
Row Labels	Sum of Profit
Bluetooth Headphones	142,887
Camcorders	5,670,098
Cameras & Camcorders Accessories	202,945
Car Video	1,243,947
Cell phones Accessories	558,835
Computers Accessories	395,441
Desktops	1,870,700
Digital Cameras	1,647,076
Digital SLR Cameras	3,854,640
Home & Office Phones	198,426
Home Theater System	2,946,634
Laptops	4,193,350
Monitors	1,268,828
Movie DVD	671,679
MP4&MP3	288,293
Printers, Scanners & Fax	1,167,342
Projectors & Screens	4,717,835
Recording Pen	160,597
Smart phones & PDAs	1,708,954
Televisions	1,388,547
Touch Screen Phones	1,173,528
VCD & DVD	118,793
Grand Total	35,589,375

Then we generate a standard Clustered Bar Chart as we have done before, and now we can also choose to use a slicer based on the Country field for the pivot table which will also dynamically affect the content of the chart which is linked to the same table.





For the next analysis, we can (using either the current pivot table or creating a new pivot table) repeat this earlier process to get the Sum of Sales and Profit by Country (Rows) and format the cells with an appropriate numeric format. Then use Product Subcategory as a slicer



2.1 Practical exercise for comparison Analysis

File to use: File to use: SalesData-v3.xlsx

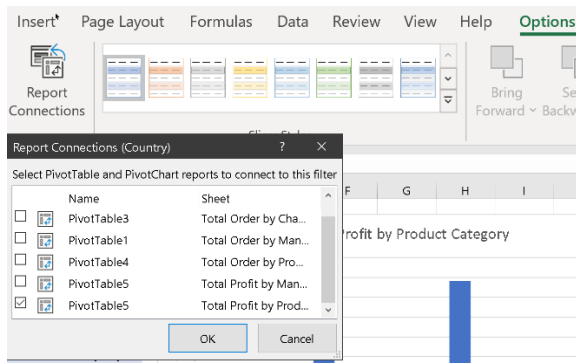
Analytical activities to perform:

1. What is the Total Order Quantity by Manufacturer?
2. What is the average Profit by Brand
3. Create a column graph displaying Total Order Quantity sold by Channel - review by Country
4. Create a bar graph displaying the Total Order quantity sold by Product Category - review by Region

Create two pivot graphs as per the instructions below. Filter both graphs by the Country and Year fields using a slicer.

5. Create a column graph displaying Total Profit by Product Category
6. Create a bar graph displaying Total Profit by Manufacturer

Note: While creating slicers, you can link a single slicer to multiple pivot tables on the same worksheet (if you decide to have more than one), using Report Connections in the Slicer Options.



3 Trend Analysis

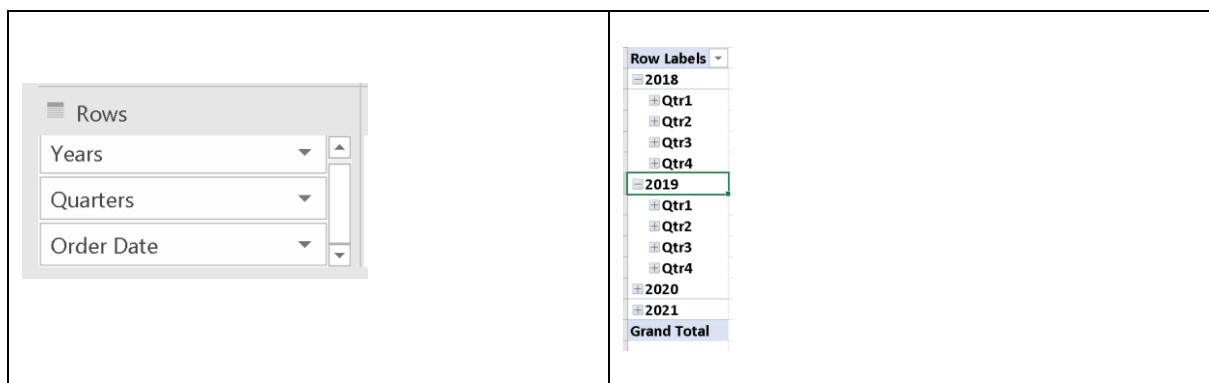
Trend analysis is all about understanding how data changes over time: is it moving up, is it moving down, is it staying stable or is it volatile. This can be accomplished mainly via trendlines. We also want to study these changes over different time frames: for e.g. changes over years, quarters, months, days. Again, this is best understood with some form of visualization.

Activities to be undertaken:

- Understand the trend for sales by year and month
- Understand the trend for sales by year and month by different product categories
- Develop a seasonality graph displaying the sales by month

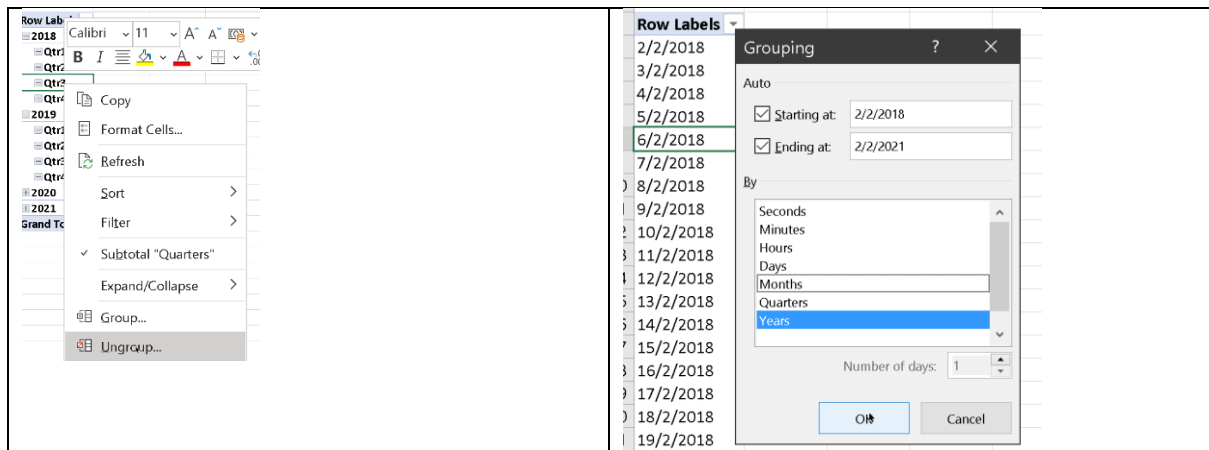
File to use: File to use: SalesData-v4.xlsx

Adding Order Date to the Rows area results in Excel automatically creating a hierarchy involving Years and Quarters.



You can ungroup the hierarchy first before creating your own custom grouping





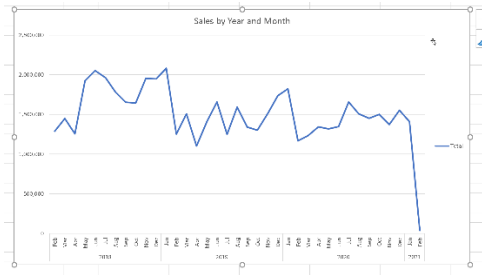
We can then continue to get Sum of Sales for each year

Row Labels	Sum of Sales
2018	18,919,151
2019	17,741,637
2020	17,284,124
2021	1,446,849
Grand Total	55,391,760

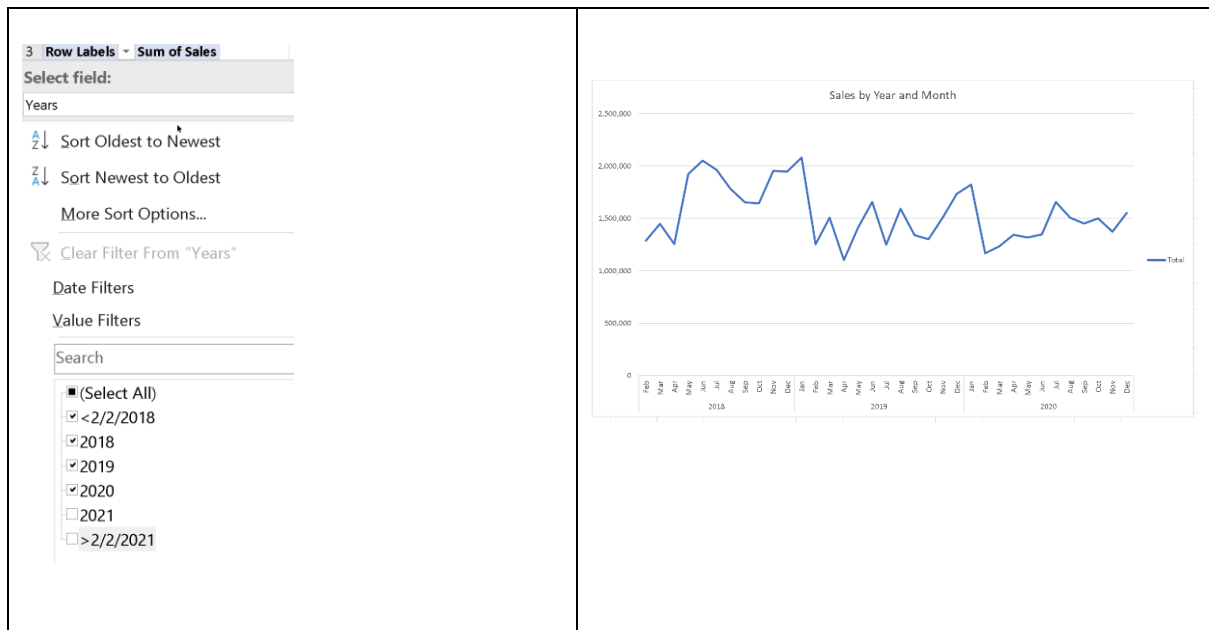
We can then alternatively change to another custom grouping (years / months)

Row Labels	Sum of Sales
2018	18,919,151
Feb	1,289,552
Mar	1,448,250
Apr	1,256,708
May	1,923,673
Jun	2,051,973
Jul	1,965,917
Aug	1,781,769
Sep	1,654,111
Oct	1,643,812
Nov	1,955,526
Dec	1,947,858
2019	17,741,637
Jan	2,081,908

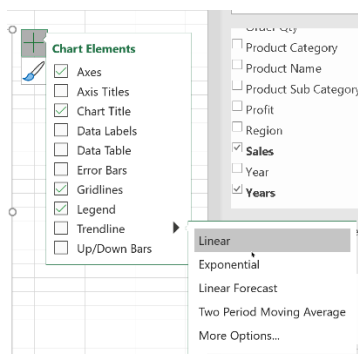
We will use a Pivot Chart to generate a Line Chart.



We can also filter out the first 2 months of 2021 to avoid issues with insufficient data at that point which will affect the trend analysis.

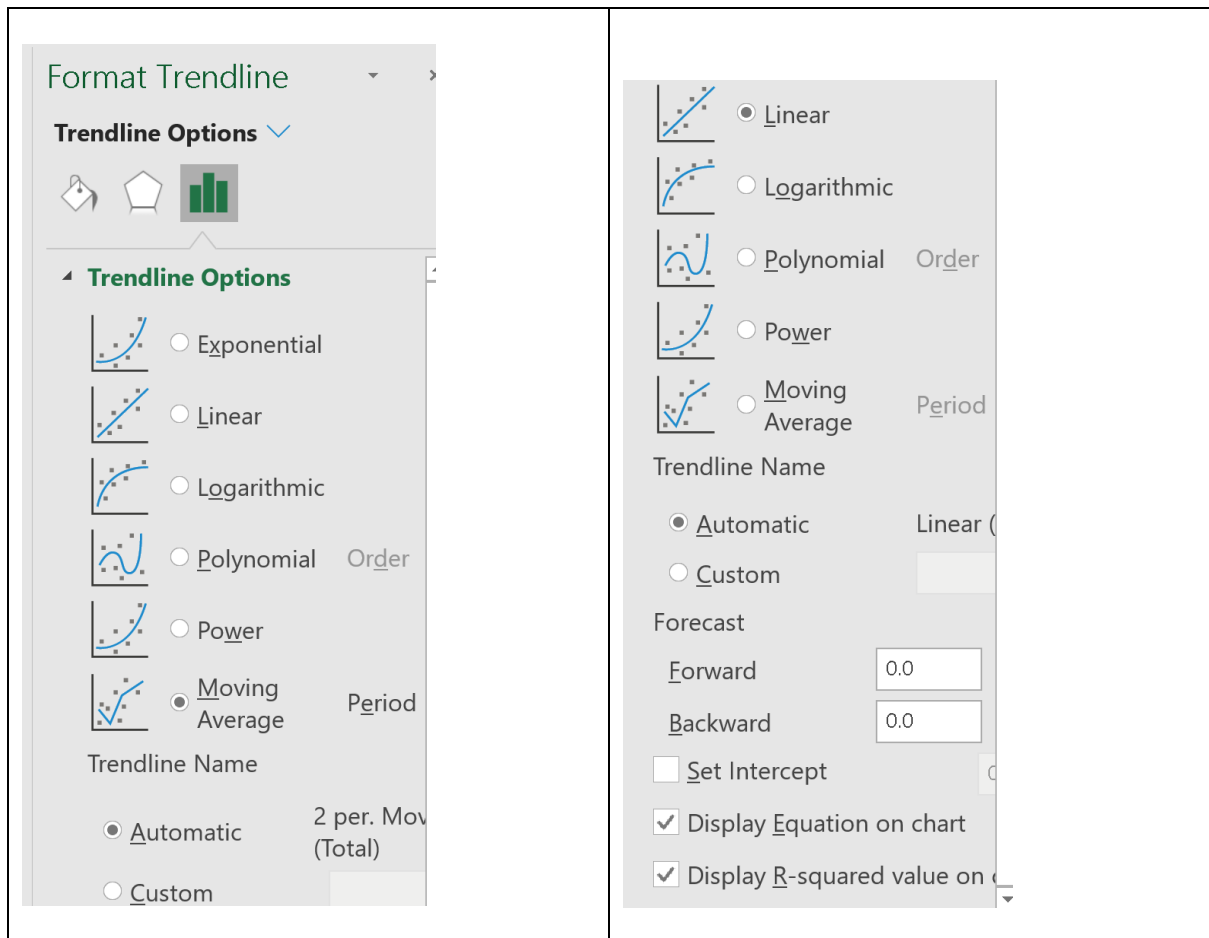


With a more proper graph, you can examine the various trend line options possible:



The first option is the default option which is linear. The rest (exponential, linear forecast) are quite similar, but not exactly the same. They are both using different algorithms to calculate what the trend line is. The two period moving average is basically taking two periods and it averages that value between the two periods. It tends to smooth the lines. Its not a trend line as such but tracking the shape of the sales lines.

In the format trend line, there are a variety of options such as polynomial, exponential, etc.



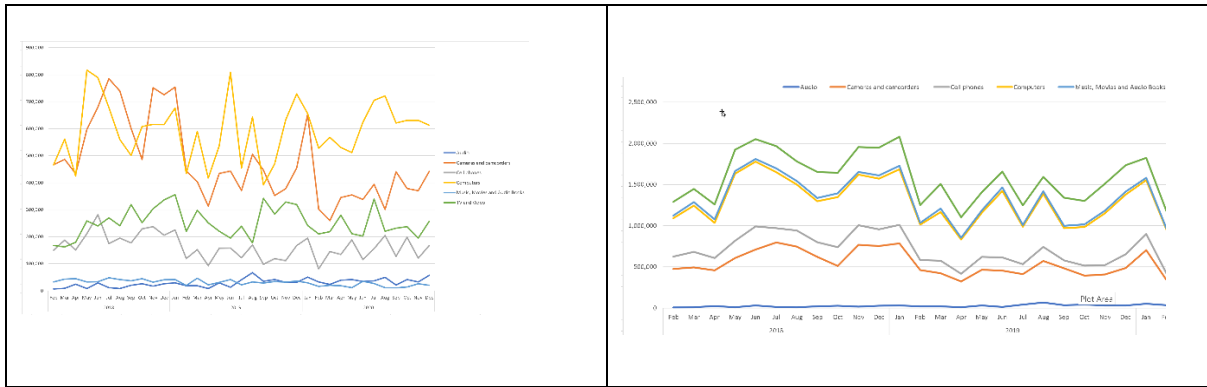
For a linear equation, you can display the trend line equation on the chart and calculate the R-squared, which is a measure of how accurate the predictions would be from the trend line equation. This is typically ranging from 0 to 1 (perfect accuracy) and is used in quantifying accuracy of multiple linear regression models.

The next step is to refine this further to understand the trend for sales by year and month by different product categories. We can simply copy the existing pivot table to a new worksheet and work from there.

Filters		Columns	
		Product Category	
Rows		Σ Values	
Years		Sum of Sales	
Order Date			

Sum of Sales		Column Labels							
Row Labels		Audio	Cameras and camcorders	Cell phones	Computers	Music	Movies and Audio Books	TV and Video	Grand Total
2018		182,424	6,750,412	2,196,698	6,635,135		426,434	2,728,048	18,919,151
Feb		6,504	466,840	149,413	466,647		32,341	167,806	1,289,552
Mar		8,254	486,074	187,716	561,066		42,948	162,191	1,448,250
Apr		24,332	433,028	150,856	424,247		44,795	179,451	1,256,708
May		7,940	597,172	210,988	816,645		32,167	258,762	1,923,673
Jun		29,136	679,810	281,364	789,193		32,827	239,643	2,051,973

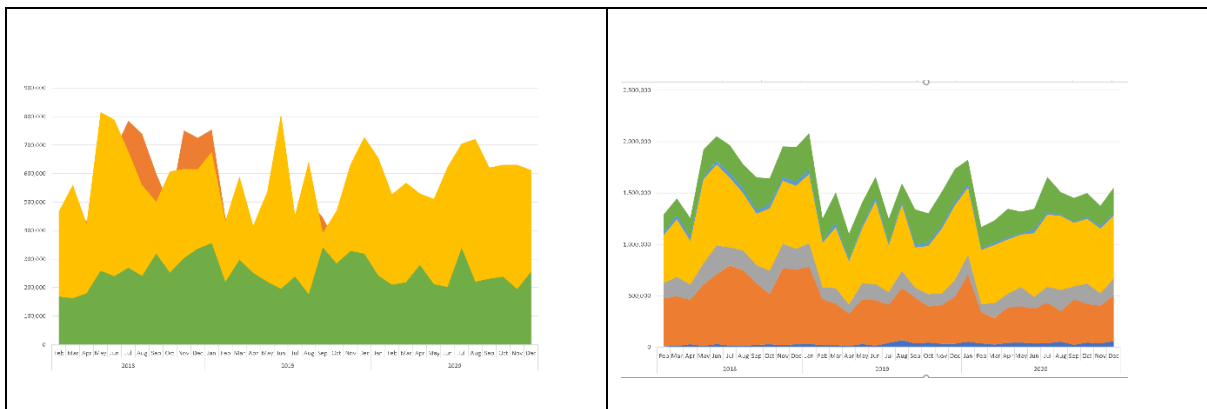
We can add a PivotChart that is a line chart and stacked line chart to get an initial view of this data.



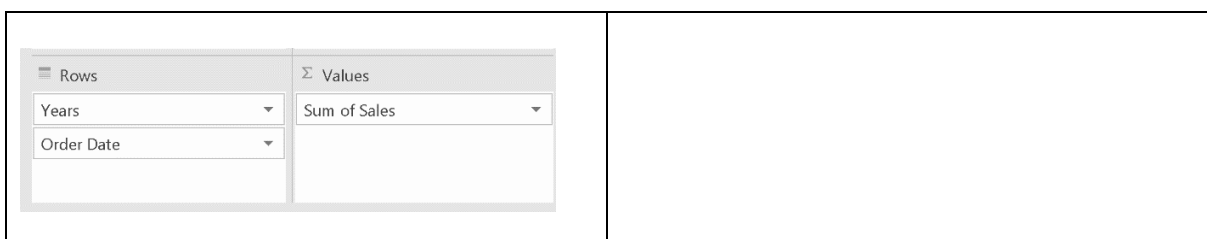
You can select which particular product category to project a trend line on if you wish.



We can also change this to either a area graph and stacked area graph to get a better view of the contributions of each product category trending over time.



You could also take your Product Category out of the Columns and utilize it as a slicer instead, which then allows you to view the trends for each of these categories across the different chart types.



Row Labels	Sum of Sales	Product Category
2018	182,424	Audio
Feb	6,504	Cameras and camcorders
Mar	8,254	Cell phones
Apr	24,332	Computers
May	7,940	Music, Movies and Audi...
Jun	29,136	TV and Video
Jul	10,904	
Aug	7,446	
Sep	19,870	
Oct	25,737	
Nov	16,790	
Dec	25,510	
2019	359,809	
Jan	29,214	
Feb	18,169	

The next thing we will look is a seasonality graph, which we can create from another copy of the pivot table on a new worksheet. For seasonality, we typically want to take into account the years which have data for all the months, and filter out those that do not.

3 Row Labels	Sum of Sales
Select field:	
Years	
Sort Oldest to Newest	
Sort Newest to Oldest	
More Sort Options...	
Clear Filter From "Years"	
Date Filters	>
Value Filters	>
Search	
(Select All)	
<2/2/2018	
2018	
2019	
2020	
2021	
>2/2/2021	

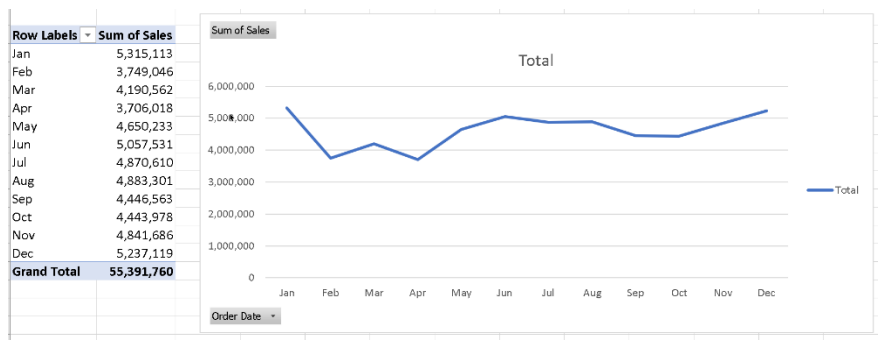
Row Labels	Sum of Sales
2019	17,741,637
Jan	2,081,908
Feb	1,252,298
Mar	1,507,751
Apr	1,103,451
May	1,408,131
Jun	1,657,980
Jul	1,248,653
Aug	1,593,694
Sep	1,340,433
Oct	1,300,579
Nov	1,511,809
Dec	1,734,950
2020	17,284,124
Jan	1,824,856
Feb	1,168,697
Mar	1,234,561
Apr	1,345,859
May	1,318,429
Jun	1,347,578
Jul	1,656,039
Aug	1,507,838
Sep	1,452,020
Oct	1,499,587
Nov	1,374,350
Dec	1,554,311
Grand Total	35,025,760

Next change custom grouping to Month. This will combine data from all the years concerned (2019, 2020).

--	--

Row Labels	Sum of Sales		Row Labels	Sum of Sales
2019			Jan	5,315,113
Jan			Feb	3,749,046
Feb			Mar	4,190,562
Mar			Apr	3,706,018
Apr			May	4,650,233
May			Jun	5,057,531
Jun			Jul	4,870,610
Jul			Aug	4,883,301
Aug			Sep	4,446,563
Sep			Oct	4,443,978
Oct			Nov	4,841,686
Nov			Dec	5,237,119
Dec			Grand Total	55,391,760
2020				
Jan				
Feb				
Mar				
Apr				
May				

Then we can generate a line graph to demonstrate the Seasonality.



Seasonality is really useful to understand so that we can understand where resource / sales demand peaks in specific periods of a year and therefore can significantly help with resource allocation planning. Its important to make sure you have a complete set of data for the entire time duration you are interested to perform analysis on (for e.g. all the months of a year, rather than partial months) because that will skew the visualization of the graph.

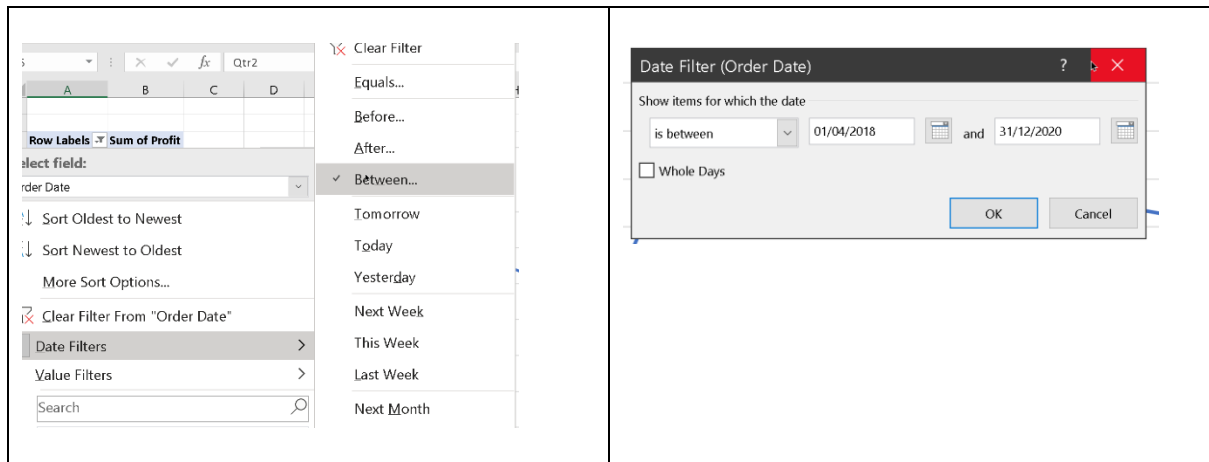
3.1 Practical exercise for Trend Analysis

Analytical activities to perform:

- Create a line graph displaying the Profit by Year and Quarter
- Add a polynomial trend line to the line graph
- Add a Slicer for Product Category
- Create an area graph displaying the Profit by Year and Quarter for the Product Category
- Add a Slicer for Country

File to use: File to use: SalesData-v5.xlsx

For first solution a), we can specifically target the quarters which are outliers in the dataset (suggesting that the data might be incomplete) and remove them using a Date filter with a Between.



Row Labels	Sum of Profit
2018	10,266,628
Qtr2	3,352,068
Qtr3	3,431,388
Qtr4	3,483,171
2019	11,378,957
Qtr1	3,172,056
Qtr2	2,541,105
Qtr3	2,731,591
Qtr4	2,934,205
2020	11,237,878
Qtr1	2,705,240
Qtr2	2,588,481
Qtr3	3,037,253
Qtr4	2,906,905
Grand Total	32,883,463

4 Ranking Analysis

Ranking analysis is about understanding the order of your items. Typically you would like the item with the highest value to be ranked number one and the item with the lowest value to have the lowest rank. So the easiest way to actually do a ranking is to sort your items.

Alternative forms of ranking including taking a selected number of items at the top and bottom of a sorted list. So you may want to see only say, the top ten items or the bottom 20 items.

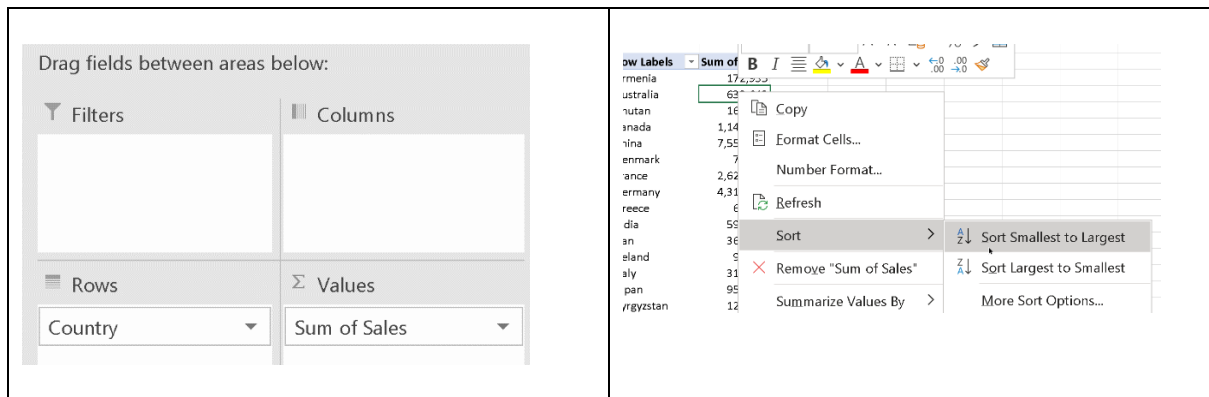
There are different ways to compute rankings and use them to affect the chronological order of how data is normally displayed (for e.g. months of a year).

Analytical activities to be performed:

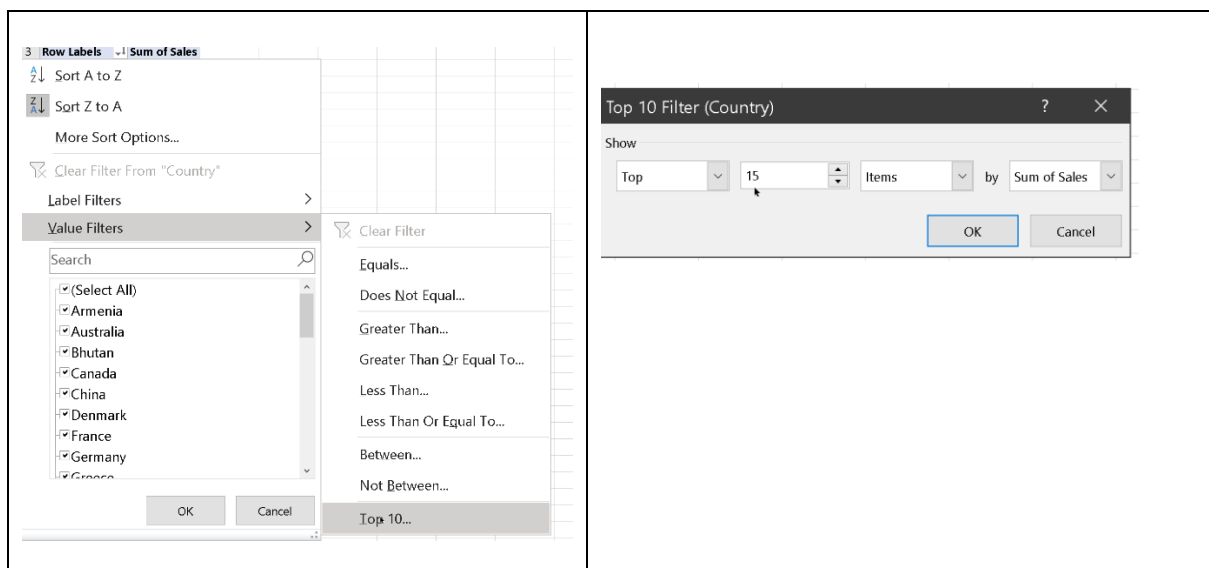
- Display top 15 countries by sales
- Display the top five product subcategories by sales
- Display the products that contribute the top 20% of profit
- Display the rankings of the profit by year and month - review by Product Category
- Display the rankings for sales and profit by different countries

File to use: SalesData-v6.xlsx

The easiest way to do a ranking is a sort of the field of interest (Sum of Sales for e.g.)



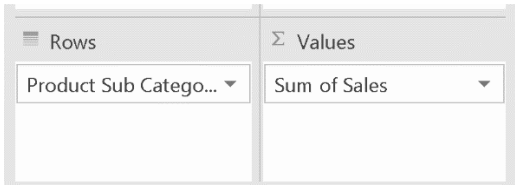
The Value Filters for a given Row Label field provides common options such as getting the top / bottom X items (rather than the entire list)

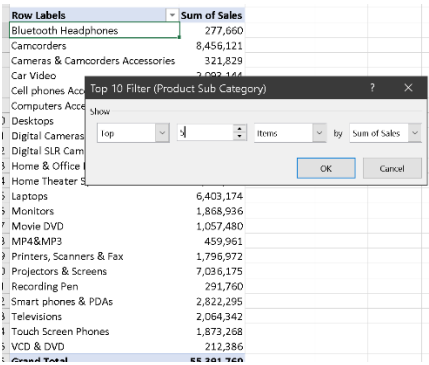


Row Labels	Sum of Sales
United States	31,635,241
China	7,557,974
Germany	4,314,956
France	2,628,401
United Kingdom	1,324,267
Canada	1,148,316
Japan	958,334
Australia	632,643
India	598,405
Russia	434,240
Turkmenistan	412,370
Iran	367,953
Syria	329,671
Italy	316,881
Pakistan	296,571
Grand Total	52,956,223

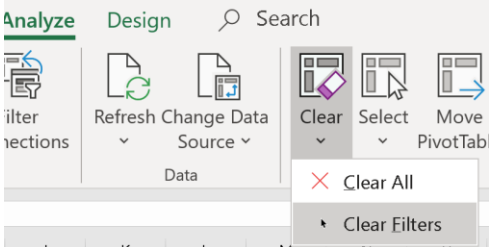
You can also further view rankings by Product Category (or any other suitable field) by adding an appropriate slicer for that field.


For the next analysis, we can repeat the use of the Value Filters for a given Row Label field to get the top 5. We can also add in another slicer for a suitable Field if we wish to.





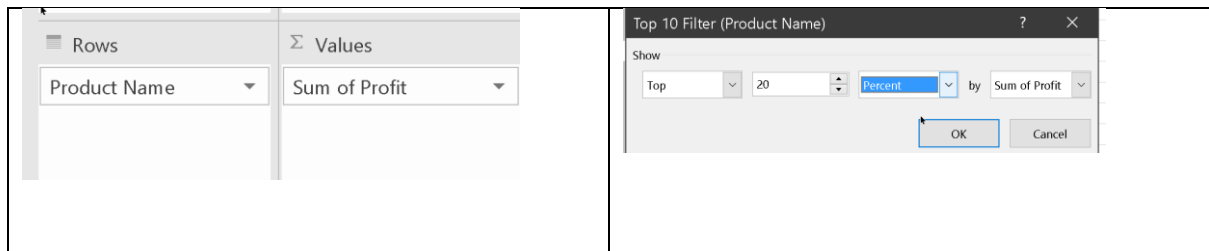
Note that you can clear the filters you have applied to the row / column labels in your Pivot table, and this is a good practice if you are going to apply multiple filters successfully.



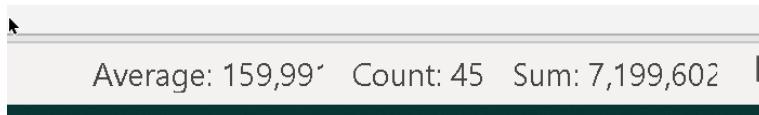


For the next analysis, we are going to filter on the top 20% instead.

--	--



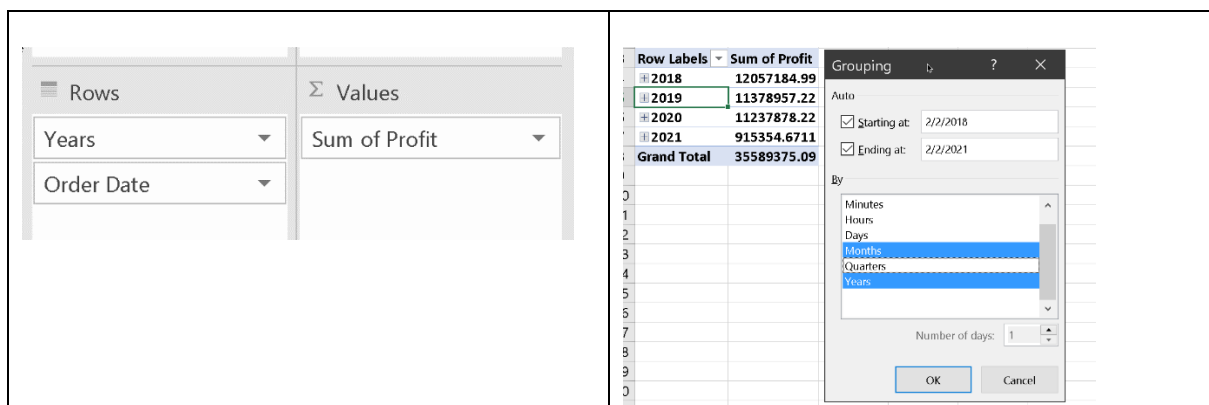
By highlighting all the values in the column for Sum of Profit, we can determine that 45 products contribute the top 20% of the profit, which is a form of Pareto Analysis.



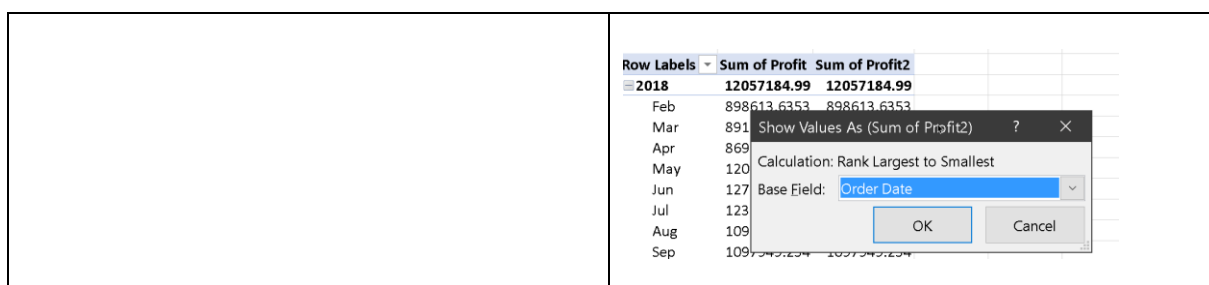
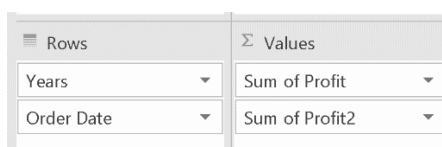
If we were to reverse this instead to find the bottom 20% of the list (in terms of overall profit), we would get slightly over 1000 products in that sub-list.

The idea would be that we would focus marketing efforts on the 45 products that contribute the top 20% of the profit, rather than all the other 1000 products

For the next analysis, we are interested in the rankings of the profit by year and month, so we need to perform the appropriate custom grouping.



To sort the months in terms of their profit total (rather than natural chronological order), we use a technique of adding in an extra sum of profit column, and performing a ranking on that instead.



Row Labels	Sum of Profit	Sum of Profit2
2018	12057184.99	12057184.99
Feb	898613.6353	898613.6353
Mar	891943.5437	891943.5437
Apr	869075.9176	869075.9176
May	1209903.159	1209903.159
Jun	1273089.103	1273089.103
Jul	1236349.967	1236349.967
Aug	1097089.286	1097089.286
Sep	1097949.234	1097949.234
Oct	1068177.876	1068177.876
Nov	1270078.95	1270078.95
Dec	1144914.316	1144914.316
2019	11378957.22	11378957.22
Jan	1324613.406	1324613.406
Feb	849994.1921	849994.1921
Mar	997448.6405	997448.6405
Apr	559962.0165	559962.0165
May	906164.75	906164.75
Jun	1074977.856	1074977.856
Jul	863281.9316	863281.9316
Aug	994217.3148	994217.3148
Sep	874091.7594	874091.7594
Oct	850178.4436	850178.4436
Nov	1001256.298	1001256.298
Dec	1082770.609	1082770.609
2020	11237878.22	11237878.22
Jan	1169711.977	1169711.977
Feb	698784.1044	698784.1044
Mar	836743.5284	836743.5284
Apr	872968.1036	872968.1036
May	871143.4547	871143.4547
Jun	888369.8336	888369.8336
Jul	1115165.676	1115165.676
Aug	988010.1517	988010.1517

Now the months are shown in correct chronological order, but we also are able to see their ranking as well (in terms of which months had the highest sales)

Row Labels	Sum of Profit	Sum of Profit2
2018	12057184.99	
Feb	898613.6353	9
Mar	891943.5437	10
Apr	869075.9176	11
May	1209903.159	4
Jun	1273089.103	1
Jul	1236349.967	3
Aug	1097089.286	7
Sep	1097949.234	6
Oct	1068177.876	8
Nov	1270078.95	2
Dec	1144914.316	5
2019	11378957.22	

The final analysis looks at the rankings for sales and profit by different countries. We can use the same technique that we did previously.

Row Labels	Sum of Sales	Sum of Sales2	Sum of Profit	Sum of Profit2
Armenia	172934.653	172934.653	135,132	135132.3122
Australia	632642.5105	632642.5105	426,288	426288.2342
Bhutan	161291.606	161291.606	116,520	116520.03
Canada	1148316.451	1148316.451	694,190	694190.0938
China	7557973.796	7557973.796	4,955,106	4955106.444
Denmark	79145.2295	79145.2295	54,735	54734.73743
France	2628400.624	2628400.624	1,500,337	1500337.194
Germany	4314956.37	4314956.37	2,672,333	2672332.508
Greece	68527.9594	68527.9594	53,280	53280.36584
India	598404.6095	598404.6095	428,186	428186.8212

The screenshot shows an Excel PivotTable with 'Country' as the row label and 'Sum of Sales' and 'Sum of Profit' as values. A context menu is open over the 'Sum of Sales' column, showing options like 'No Calculation', '% of Grand Total', etc. The 'Show Values As' option is selected, opening a dialog box. The dialog box shows 'Calculation: Rank Largest to Smallest' and 'Base Field: Country'.

Row Labels	Sum of Sales	Sum of Sales2	Sum of Profit	Sum of Profit2
Armenia	172934.653	172934.653	135,132	135132.3122
Australia	632642.5105	632642.5105	426,288	426288.2342
Bhutan	161291.606	161291.606	116,520	116520.03
Canada	1148316.451	1148316.451	694,190	694190.938
China	7557973.796	7557973.796	4,955,106	4955106.444
Denmark	79145.2295	79145.2295	54,735	54735.743
France	2628400.624	2628400.624	1,500,337	1500337.194
Germany	4314956.37	4314956.37	2,672,333	2672333.508
Greece	68527.9594	68527.9594	53,280	53280.584
India	598404.6095	598404.6095	428,186	428186.8212

We can repeat this again for the sum of Profit to get rankings for both Sales and Profit, which allows use to comparatively compare across these 2 categories. This may be useful for us to identify anomalies where the sales ranking might be very high but the profit ranking is much lower or vice versa.

Row Labels	Total Sales	Sales Ranking	Total Profit	Profit Ranking
Armenia	172934.653	19	135,132	18
Australia	632642.5105	8	426,288	9
Bhutan	161291.606	20	116,520	20
Canada	1148316.451	6	694,190	6
China	7557973.796	2	4,955,106	2
Denmark	79145.2295	32	54,735	33
France	2628400.624	4	1,500,337	4
Germany	4314956.37	3	2,672,333	3
Greece	68527.9594	34	53,280	34
India	598404.6095	9	428,186	8
Iran	367953.336	12	193,023	16

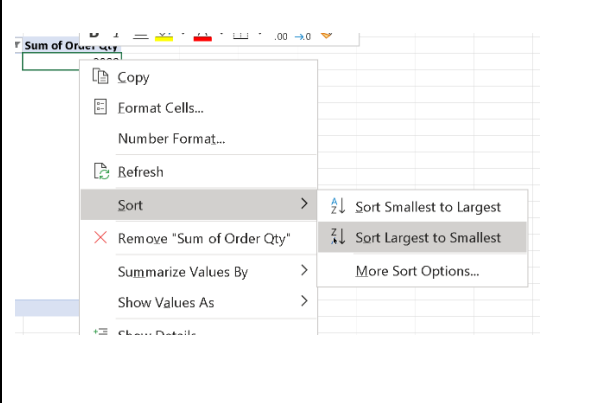
4.1 Practical Exercise for Ranking Analysis

Analytical activities to perform:

- Identify the Top 15 Products by Order Quantity sold, display the total Sales and Profit and add a Slicer for Country
- Create a Ranking for Product Name based on Total Sales and Profit, showing the ranking for both Sales and Profit separately in 2 columns

File to use: SalesData-v7.xlsx

For a), note that after the initial filter for top 15, you can further perform a sort to get the sorted according to sum of order quantity, since the default sorting order is the values in the row label (which are alphabetical values)

	<table border="1"> <thead> <tr> <th>Row Labels</th><th>Sum of Order Qty</th></tr> </thead> <tbody> <tr><td>Contoso In-Line Coupler E180 Silver</td><td>6780</td></tr> <tr><td>Headphone Adapter for Contoso Phone E130 Silver</td><td>4680</td></tr> <tr><td>Contoso Rubberized Snap-On Cover Hard Case Cell Phone Protector E160 Silver</td><td>4580</td></tr> <tr><td>Contoso In-Line Coupler E180 White</td><td>4060</td></tr> <tr><td>Contoso Rubberized Skin BlackBerry E100 Silver</td><td>3560</td></tr> <tr><td>Cigarette Lighter Adapter for Contoso Phones E110 White</td><td>3520</td></tr> <tr><td>Contoso Rubberized Snap-On Cover Hard Case Cell Phone Protector E160 Pink</td><td>3280</td></tr> <tr><td>Contoso Touch Stylus Pen E150 Black</td><td>3060</td></tr> <tr><td>Cigarette Lighter Adapter for Contoso Phones E110 Red</td><td>2900</td></tr> <tr><td>Contoso Original K1m Li-Ion Standard Battery E170 Black</td><td>2860</td></tr> <tr><td>Contoso In-Line Coupler E180 Black</td><td>2720</td></tr> <tr><td>Contoso Rubberized Snap-On Cover Hard Case Cell Phone Protector E160 White</td><td>2700</td></tr> <tr><td>Contoso Touch Stylus Pen E150 Red</td><td>2600</td></tr> <tr><td>Contoso Rubberized Skin BlackBerry E100 Black</td><td>2560</td></tr> <tr><td>Contoso Original K1m Li-Ion Standard Battery E170 Silver</td><td>2480</td></tr> <tr><td>Grand Total</td><td>52340</td></tr> </tbody> </table>	Row Labels	Sum of Order Qty	Contoso In-Line Coupler E180 Silver	6780	Headphone Adapter for Contoso Phone E130 Silver	4680	Contoso Rubberized Snap-On Cover Hard Case Cell Phone Protector E160 Silver	4580	Contoso In-Line Coupler E180 White	4060	Contoso Rubberized Skin BlackBerry E100 Silver	3560	Cigarette Lighter Adapter for Contoso Phones E110 White	3520	Contoso Rubberized Snap-On Cover Hard Case Cell Phone Protector E160 Pink	3280	Contoso Touch Stylus Pen E150 Black	3060	Cigarette Lighter Adapter for Contoso Phones E110 Red	2900	Contoso Original K1m Li-Ion Standard Battery E170 Black	2860	Contoso In-Line Coupler E180 Black	2720	Contoso Rubberized Snap-On Cover Hard Case Cell Phone Protector E160 White	2700	Contoso Touch Stylus Pen E150 Red	2600	Contoso Rubberized Skin BlackBerry E100 Black	2560	Contoso Original K1m Li-Ion Standard Battery E170 Silver	2480	Grand Total	52340
Row Labels	Sum of Order Qty																																		
Contoso In-Line Coupler E180 Silver	6780																																		
Headphone Adapter for Contoso Phone E130 Silver	4680																																		
Contoso Rubberized Snap-On Cover Hard Case Cell Phone Protector E160 Silver	4580																																		
Contoso In-Line Coupler E180 White	4060																																		
Contoso Rubberized Skin BlackBerry E100 Silver	3560																																		
Cigarette Lighter Adapter for Contoso Phones E110 White	3520																																		
Contoso Rubberized Snap-On Cover Hard Case Cell Phone Protector E160 Pink	3280																																		
Contoso Touch Stylus Pen E150 Black	3060																																		
Cigarette Lighter Adapter for Contoso Phones E110 Red	2900																																		
Contoso Original K1m Li-Ion Standard Battery E170 Black	2860																																		
Contoso In-Line Coupler E180 Black	2720																																		
Contoso Rubberized Snap-On Cover Hard Case Cell Phone Protector E160 White	2700																																		
Contoso Touch Stylus Pen E150 Red	2600																																		
Contoso Rubberized Skin BlackBerry E100 Black	2560																																		
Contoso Original K1m Li-Ion Standard Battery E170 Silver	2480																																		
Grand Total	52340																																		

For b) , you can also sort again on the separate ranking columns if you wish for Ranking on Sales and Ranking on Profit. This allows you to see correlation between ranking on sales and profit for various items.