Data Analysis Utilizing Excel - Part 1
By Palani Murugappan

Understanding data analysis

Data analysis is a process of seeking meaningful information from raw data and suggesting conclusions from the analysis done by using some tools such as Microsoft Excel.

Why use Microsoft Excel? It has been shown to be widely popular and one of the more commonly used spreadsheet software that has a tremendous variety of built-in functions and features that can be tapped into the areas of informed decision making.

Selecting lowest two quotations from various vendors

Let us look at a scenario as below. You are given raw data of various Vendors who had quoted their prices on a weekly basis from 5 Sept onwards.

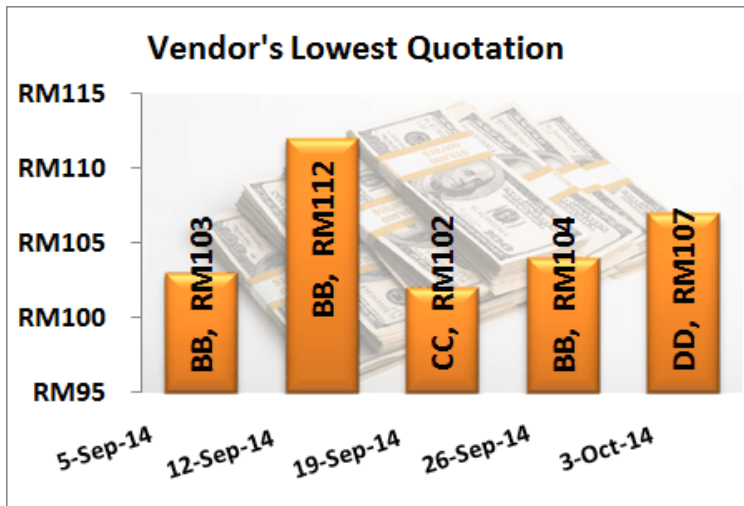| Vendor	AA	BB	CC	DD
5-Sep-14	RM 147	RM 103	RM 119	RM 149
12-Sep-14	RM 121	RM 112	RM 130	RM 137
19-Sep-14	RM 131	RM 121	RM 102	RM 128
26-Sep-14	RM 146	RM 104	RM 118	RM 134
3-Oct-14	RM 127	RM 121	RM 136	RM 107

How will you determine the two lowest quotations from above? By looking at the data, the lowest quotation for each week is highlighted in red. This process is automated using a feature of Excel called Conditional Formatting.

Next, how do you determine the second lowest quote from above and how do you automate the process such that if same Vendor decides to have a change of heart and reduce the quote?
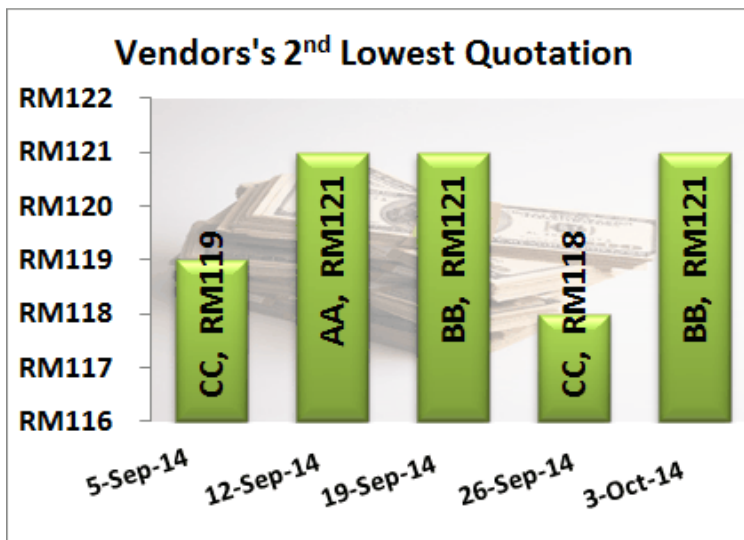
Using functions of Excel, one can automate the whole selection process of the lowest and second lowest quotation automatically where if the user updates the figures from Vendors, the listing below will scan and select the lowest two figures automatically all the time.

Vendor	Lowest		Vendor	2nd Lowest	
BB	RM 103		CC	RM	119
BB	RM 112		AA	RM	121
CC	RM 102		BB	RM	121
BB	RM 104		CC	RM	118
DD	RM 107		BB	RM	121

The plot of the above data with the lowest quotation is displayed below and again, this plot of column chart will update itself automatically when the above figures changes.

The second lowest quotation is also plotted beside the above column chart to show a similar scenario. Note that it is quite possible for the same Vendor to quote the lowest and second lowest prices on a weekly basis.

Selection of Vendor is then based on criteria set by the purchasing department.

Analyzing sales data

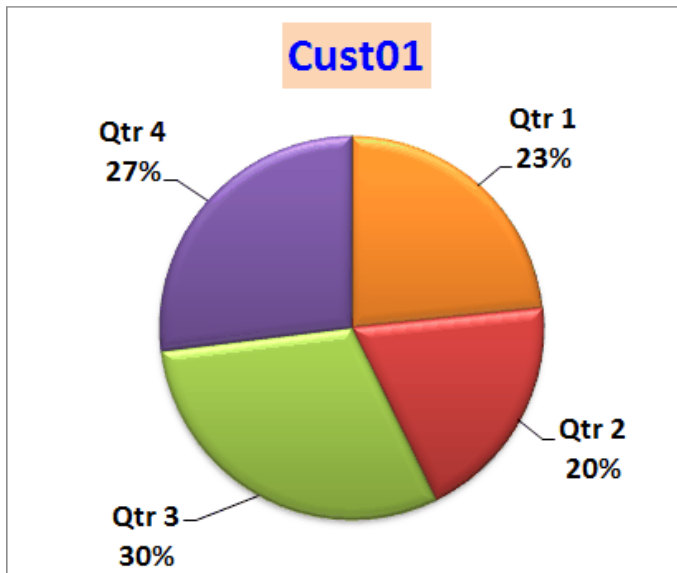In the next scenario, one is given the following sales data as below.

Customer	Region	Qtr 1	Qtr 2	Qtr 3	Qtr 4
Cust01	North	259	215	336	299
Cust02	South	458	440	252	382
Cust03	East	174	393	382	109
Cust04	West	145	439	268	109
Cust05	North	181	262	248	136
Cust06	South	288	484	193	233
Cust07	East	321	318	150	460
Cust08	West	242	108	190	272
Cust09	North	142	159	218	106
Cust10	South	180	141	195	334
Cust11	East	152	477	453	194
Cust12	West	394	168	379	295

What information can we gather from the data above? In a real world situation, there are softwares to perform some simple analysis. However, if one is familiar with Excel's functions and features, one can extract very specific information from the sales data above. Let us look at what information can be gathered.

For a start, the data above can be summarized on different cells to display a specific customer and the sales per quarter as displayed below. We can gather the maximum, minimum, and average value of the 4 quarters. In the screen capture below, one can see that the maximum sales from Cust01 is in the 3rd quarter with a figure of 336 while the minimum figure from Cust01 is 215 which took place in the 2nd quarter.

	Qtr 1	Qtr 2	Qtr 3	Qtr 4
Cust01	259	215	336	299
	MAX	MIN	AVERAGE	
Cust01	336	215	277	
Quarter	Qtr 3	Qtr 2		

The above data is linked to a pie chart that changes automatically based on selection of customer too.
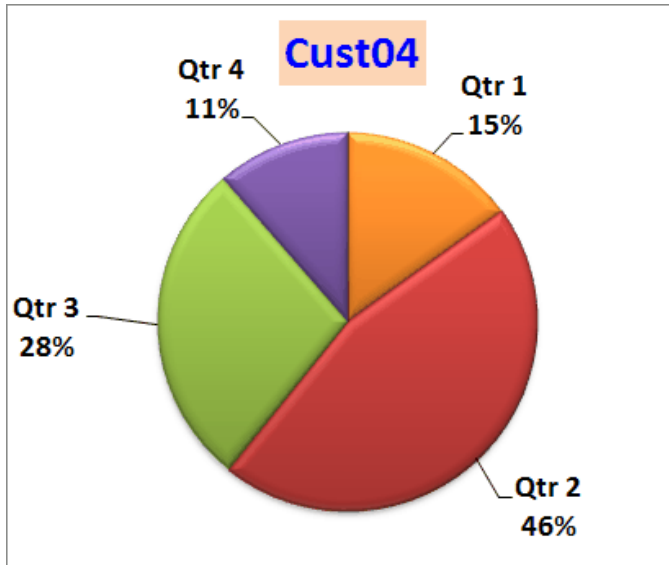
Now, let us change the selection of customer to Cust04 form the drop down list of options.

15		Qtr 1	Qtr 2	Qtr 3	Qtr 4
16 **Cust01**	▾	259	215	336	299

Cust01	
Cust02	
Cust03	AX	MIN	AVERAGE
Cust04	
Cust05	336	215	277
Cust06	Qtr 3	Qtr 2
Cust07	
Cust08	

Upon selecting Cust04, the maximum, minimum, and average value updates automatically as in the screen capture below.

		Qtr 1	Qtr 2	Qtr 3	Qtr 4
Cust04	▾	145	439	268	109

	MAX	MIN	AVERAGE
Cust04	439	109	240
Quarter	Qtr 2	Qtr 4	

As can be seen, Excel is a fantastic tool to perform "what-if" analysis by choosing different customers from the drop down list of options. Again, the chart updates itself when a new customer is selected from the drop down list.

What other analysis can we gather from the sales data above? We can specify different criteria in different cells within the spreadsheet as follows.

Criteria	
Region	**Qtr 1**
North	**>200**
Average	259
Count	1
Total Sum	259
Max	259
Min	259

In the above example, if the user selected the North Region with the criteria of sales greater than 200 for Quarter 1, the database functions of Excel can summarize and inform us that there is only one sales of value above 200.

However, if we change the region to West from the drop down list option, the figures are totally different as below.

For the above West region, one can see that there are two instances when sales had exceeded 200 in Quarter 1 with the maximum value of 394 and minimum value of 242. If one is not familiar with Excel functions, the above analysis has to done manually by sorting the data and grouping the same categories of data together. This can be very cumbersome!

Going a step further, one has the option of changing the quarter to another. In the illustration below, Quarter 1 has been changed to 4 and the figures within the criteria updates automatically as can be seen below.

Criteria	
Region	**Qtr 4**
West _(Qtr 1 / Qtr 2 / Qtr 3 / **Qtr 4**)_	
Average	
Count	2
Total Sum	567
Max	295
Min	272

→

Criteria	
Region	**Qtr 4**
West	**>200**
Average	284
Count	2
Total Sum	567
Max	295
Min	272

Finally, just like the Region and Quarters, the use also can specify different values for the quarters. In the above illustration, we had selected values greater than 200. By clicking on the drop down list option of this cell, the user has other option to choose from to (the options listed here are all user defined). If the user changes the option from ">200" to "<150", the figures updates automatically yet once again as follows. As can be seen, there is only one occurrence of values less than 150 for the 4th quarter of the West region.

Criteria	
Region	**Qtr 4**
West	**>200** _(<150 / <250 / >400 / >300 / >200)_
Average	
Count	
Total Sum	567
Max	295
Min	272

→

Criteria	
Region	**Qtr 4**
West	**<150**
Average	109
Count	1
Total Sum	109
Max	109
Min	109

The above was a simple illustration of the type of analysis one can do to gather meaningful information from raw data using tools such as Microsoft Excel. Stay tuned for the second part of data analysis…..

The examples illustrated above are a snapshot of what the author covers in some of his 2-days training sessions. Having been a corporate trainer for the last 14 years, Palani specializes in training participants how best to perform data analysis from company's raw data and concluding from the information gathered by harnessing Excel's built-in functions and features. A HRDF certified trainer and author of 51 books to-date on various software applications, Palani is passionate on what Excel can do, and how best to present data in different formats. More information is covered in his website www.avimursolutions.com. He can be reached at palani@avimursolutions.com