Data Analysis Utilizing Excel - Part 2

By Palani Murugappan

Understanding data analysis

In the last article, the process of data analysis in terms of selecting the lowest two quotations from various vendors; and analyzing sales data were shown. In this article, we will probe the data further to gather other meaningful information.

Analyzing sales data using Excel functions

Let us assume you have the following sales figures by Region for the various customers. Figures are compilation of total sales amount on a quarterly basis.

	A	B	C	D	E	F
1	**Customer**	**Region**	**Qtr 1**	**Qtr 2**	**Qtr 3**	**Qtr 4**
2	**Cust01**	North	231	330	364	456
3	**Cust02**	South	331	184	272	381
4	**Cust03**	East		335	354	335
5	**Cust04**	West	151	424	196	187
6	**Cust05**	North	403	412	324	363
7	**Cust06**	South			402	253
8	**Cust07**	East	354	415	204	343
9	**Cust08**	West	317	404	497	243
10	**Cust09**	North	203	461	267	486
11	**Cust10**	South	498	178	233	320
12	**Cust11**	East				379
13	**Cust12**	West	338	152	249	333

Based on the above data, let us look at what type of analysis that can be done. For example, when a sales manager looks at the above data, he/she may want to look at how many blank cells (representing zero sales) are there. This can be achieved by using the COUNTBLANK function in Excel with the following results:

Sales Data Analysis

1) Total number of blank results for the various Quarters:

Quarter 1:	**3**
Quarter 2:	**2**
Quarter 3:	**1**
Quarter 4:	**0**

The results above indicates that there are 3, 2, 1, and 0 blank cells based on the Quarters 1, 2, 3, and 4 respectively.

Next, expanding on the COUNTBLANK function, how would the sales manager determine which Regions have zero sales based on above? Here, the Excel function COUNTIFS is used to determine the total number of blank cells based on Regions as follows:

2) Total number of blank results based on various Regions:

North	**0**
South	**2**
East	**4**
West	**0**

Based on the above data, the sales manager may now want to do an analysis by Region and by Quarter. This can be summarized and simplified as follows:

3) Total number of blank results based on Quarters and Regions:

Region	Qtr1	Qtr2	Qtr3	Qtr4	Total
North	**0**	**0**	**0**	**0**	**0**
South	**1**	**1**	**0**	**0**	**2**
East	**2**	**1**	**1**	**0**	**4**
West	**0**	**0**	**0**	**0**	**0**

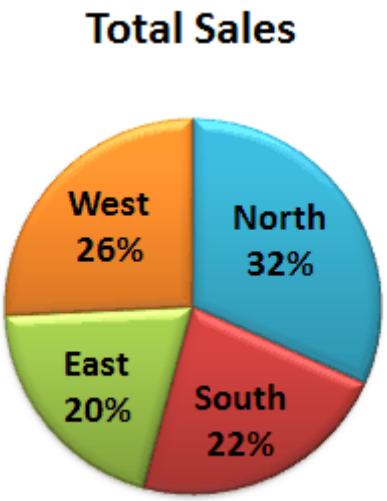
The Total column results matches that in point 2 calculated previously. Note that all calculations are performed using built-in functions of Excel; nothing is done by manual selection of cells.

Next, let us perform the calculation of Total Sales by Region. Here, the Excel function SUMIF is used to select a specific region from all of the regions and perform the calculations accordingly. The tabulated Total Sales are as follows:

4) Total Sales by Regions

Region	Total Sales
North	**4,300**
South	**3,052**
East	**2,719**
West	**3,491**

Based on the above Total Sales figures, a pie chart is created to depict the segments based on percentage of sales amount as follows.

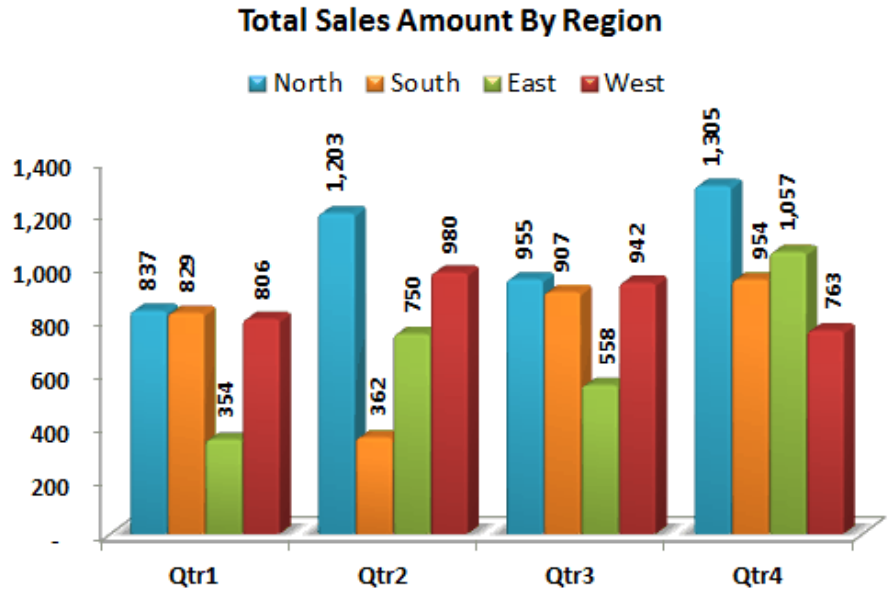


Moving on, by probing the Total Sales more in detail, a sample table by Region and Quarters can be created as follows. The calculations are based on using the SUMIF function once again.

5) Total Sales Amount by Regions and Quarters:

Region	Qtr1	Qtr2	Qtr3	Qtr4	Total
North	837	1,203	955	1,305	4,300
South	829	362	907	954	3,052
East	354	750	558	1,057	2,719
West	806	980	942	763	3,491

The above data can be represented by a 3-D column chart as below:

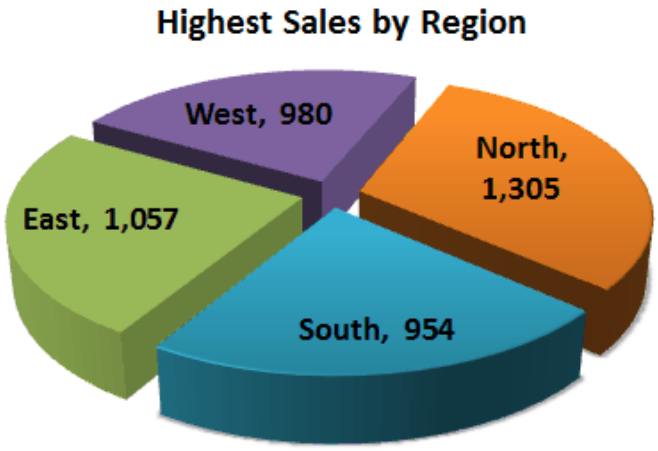


The data above can be summarized further by analyzing which Quarter produced the highest Sales as follows. This is done utilizing the Nested IF function in Excel together with the function MAX.

6) Highest Sales by Region & Quarter:

	North	**1,305**	**Qtr4**
	South	**954**	**Qtr4**
	East	**1,057**	**Qtr4**
	West	**980**	**Qtr2**

Once again, the data of highest sales can be depicted using an exploded pie chart as illustrated below.



As in the previous article, the feature called Conditional Formatting is used to highlight the highest sales by Region in red as illustrated.

Region	Qtr1	Qtr2	Qtr3	Qtr4	Total
North	**837**	**1,203**	**955**	**1,305**	**4,300**
South	**829**	**362**	**907**	**954**	**3,052**
East	**354**	**750**	**558**	**1,057**	**2,719**
West	**806**	**980**	**942**	**763**	**3,491**

What can one conclude from the above? For the general person, it seems obvious that the salesmen are going “all out” to gather the most amount of sales in the final Quarter 4 (with the exception of Quarter 2).

In the next example, you shall be shown a “different” aspect of data analysis by using the feature called Pivot Table.

Analyzing sales data using Pivot Table

Another quicker way of gathering information by performing data analysis is using the Pivot Table. A pivot table can be invoked by highlighting the entire data (with the header) and clicking on the Pivot Table command button within the Insert tab. By dragging and dropping the relevant fields in specific areas, a summary of the above data is represented as below.

Row Labels	Sum of Qtr 1	Sum of Qtr 2	Sum of Qtr 3	Sum of Qtr 4
East	354	750	558	1,057
North	837	1,203	955	1,305
South	829	362	907	954
West	806	980	942	763
Grand Total	**2,826**	**3,295**	**3,362**	**4,079**

Note that the Region is presented based on ascending order. This is the default option when pivot table is invoked. The first column of data is presented in an ascending manner.

To give the pivot table a fancy look, one can apply the Conditional Formatting Data Bars option as illustrated.

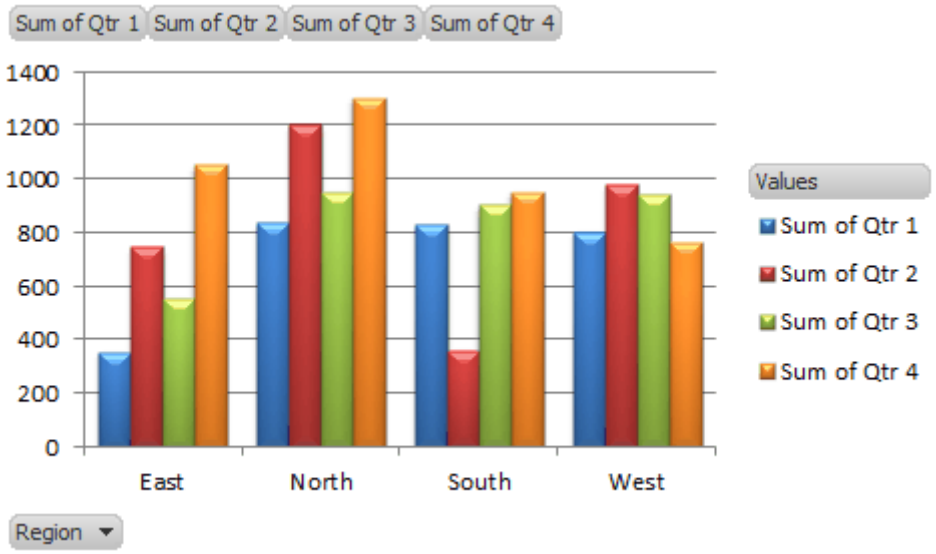
Row Labels	Sum of Qtr 1	Sum of Qtr 2	Sum of Qtr 3	Sum of Qtr 4
East	354	750	558	1,057
North	837	1,203	955	1,305
South	829	362	907	954
West	806	980	942	763
Grand Total	**2,826**	**3,295**	**3,362**	**4,079**

Besides the option of Data bars, the user can also select the Icon Sets option within Conditional Formatting. An example of Icon Set is applied and displayed below.

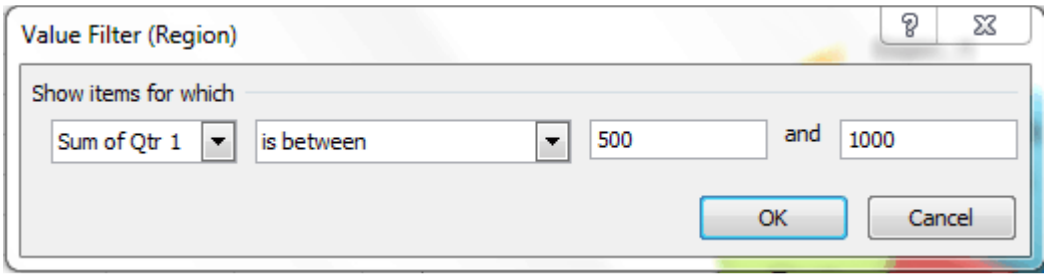
Row Labels	Sum of Qtr 1	Sum of Qtr 2	Sum of Qtr 3	Sum of Qtr 4
East	354	750	558	1,057
North	837	1,203	955	1,305
South	829	362	907	954
West	806	980	942	763
Grand Total	**2,826**	**3,295**	**3,362**	**4,079**

The Icon Sets can be set based specific conditions where the three colors i.e. green, yellow and red are displayed based on conditions satisfied e.g. Sales Amount greater than 1000 in green.

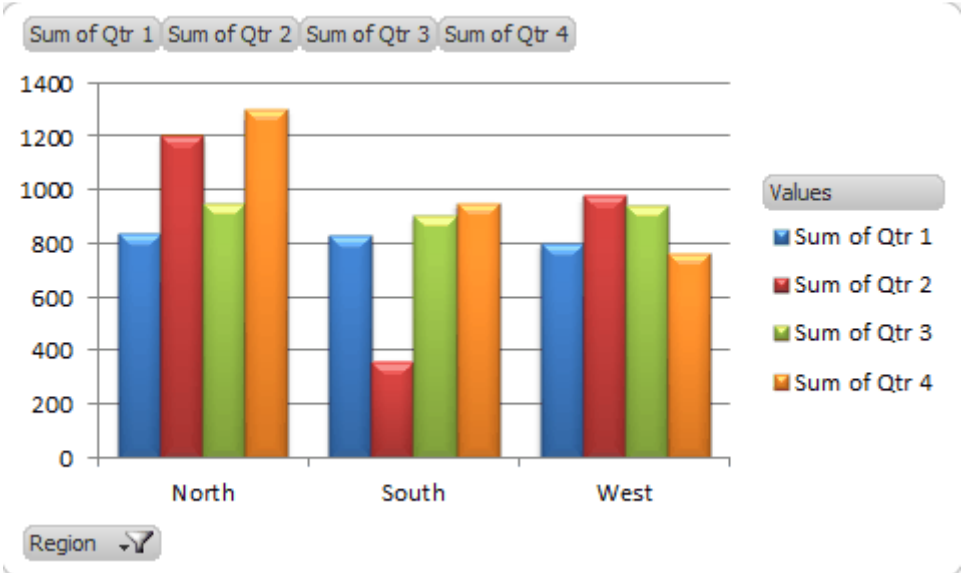
Based on the compiled data within the Pivot Table, a Pivot Chart can be created as in the illustration.



Note that based on the chart above, one can filter the chart to display values between 500 to 1000 only for Quarter 1 as in the dialog box option below.



This then alters the appearance of the chart to display the columns in Quarter 1 with values between selected range only.



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