Data Mining and Analysis With Excel PivotTables and QI Macros

By Jay Arthur, The KnowWare® Man

QIMacros®

Data Mining and Analysis
with Excel PivotTables and QI Macros

Table of Contents

Data Mining and Analysis
with Excel PivotTables and QI Macros
by Jay Arthur, The KnowWare® Man

What Gets Measured Gets Done

It's an old, but true saying that what gets measured gets done. That's why so many companies are looking to measure their performance in ways that drive optimal performance.

This report will show you:
1. How to use Pivot Tables to mine your data
2. How to use the QI Macros to chart your data
3. Other Tips and Tricks

If you measure the wrong thing, you get wrong-headed behavior. If you measure the right thing, you get right behavior. It's that simple. So, before you start setting up a pivot table, you need to make sure you're measuring the things that will drive your business forward, not backward.

To succeed at Six Sigma or any process improvement effort, you'll often have to analyze and summarize text data. Most companies have lots of transaction data from "flat files" like the one shown below, but because the data consists of text and raw numbers, they sometimes have a hard time figuring out what to do with it.

	A	B	C	D	E	F	G	H	I	J	K	L
1	DRG	Physi-cian ID #	DRG	APS DRG	Diagnosis	Age	Sex	LOS	Total Charge	Date	Discharge Status	Adverse Event(s)
2	373: Vaginal Delivery	MD11	373	3730	664.01: First-Degree Peri	26	F	2	$ 5,729	10/1/2006	Home, Self-Care	--
3	373: Vaginal Delivery	MD11	373	3730	645.11: Post Term Pregnan	18	F	2	$ 9,551	10/2/2006	Home, Self-Care	--
4	373: Vaginal Delivery	MD8	373	3730	663.31: Oth Unspec Cord E	37	F	1	$ 6,976	10/2/2006	Home, Self-Care	--
5	373: Vaginal Delivery	MD8	373	3730	650: Normal Delivery	19	F	1	$ 4,589	10/3/2006	Home, Self-Care	--
6	373: Vaginal Delivery	MD1	373	3730	650: Normal Delivery	28	F	2	$ 11,033	10/4/2006	Home, Self-Care	--
7	373: Vaginal Delivery	MD2	373	3730	663.31: Oth Unspec Cord E	27	F	1	$ 7,002	10/4/2006	Home, Self-Care	--
8	373: Vaginal Delivery	MD3	373	3730	646.81: Oth Spec Complica	24	F	2	$ 7,190	10/4/2006	Home, Self-Care	--
9	373: Vaginal Delivery	MD3	373	3730	645.11: Post Term Pregnan	21	F	2	$ 6,313	10/4/2006	Home, Self-Care	--
10	373: Vaginal Delivery	MD5	373	3730	650: Normal Delivery	19	F	1	$ 6,377	10/4/2006	Home, Self-Care	--
11	373: Vaginal Delivery	MD11	373	3730	656.61: Excessive Fetal G	22	F	1	$ 7,778	10/5/2006	Home, Self-Care	--
12	373: Vaginal Delivery	MD3	373	3730	664.01: First-Degree Peri	19	F	1	$ 6,753	10/5/2006	Home, Self-Care	--
13	373: Vaginal Delivery	MD6	373	3730	663.31: Oth Unspec Cord E	22	F	1	$ 8,369	10/5/2006	Home, Self-Care	Complication
14	373: Vaginal Delivery	MD6	373	3730	664.11: Second-Degree Per	24	F	2	$ 7,292	10/5/2006	Home, Self-Care	--
15	373: Vaginal Delivery	MD2	373	3730	664.01: First-Degree Peri	26	F	1	$ 6,464	10/6/2006	Home, Self-Care	Complication
16	373: Vaginal Delivery	MD3	373	3730	664.11: Second-Degree Per	20	F	1	$ 6,425	10/6/2006	Home, Self-Care	--
17	373: Vaginal Delivery	MD5	373	3730	648.91: Oth Current Condi	28	F	2	$ 7,985	10/6/2006	Home, Self-Care	--
18	373: Vaginal Delivery	MD8	373	3730	650: Normal Delivery	19	F	1	$ 7,299	10/6/2006	Home, Self-Care	--
19	373: Vaginal Delivery	MD8	373	3730	645.11: Post Term Pregnan	19	F	2	$ 8,344	10/7/2006	Home, Self-Care	--
20	373: Vaginal Delivery	MD2	373	3730	656.11: Rhesus Isoimmuniz	33	F	1	$ 7,880	10/8/2006	Home, Self-Care	--
21	373: Vaginal Delivery	MD2	373	3730	645.11: Post Term Pregnan	32	F	2	$ 9,488	10/8/2006	Home, Self-Care	--
22	373: Vaginal Delivery	MD6	373	3730	648.91: Oth Current Condi	35	F	2	$ 7,521	10/8/2006	Home, Self-Care	--
23	373: Vaginal Delivery	MD6	373	3730	664.01: First-Degree Peri	21	F	2	$ 5,298	10/8/2006	Home, Self-Care	--
24	373: Vaginal Delivery	MD6	373	3730	664.11: Second-Degree Per	15	F	1	$ 6,514	10/8/2006	Home, Self-Care	--
25	373: Vaginal Delivery	MD6	373	3730	663.11: Cord Around Neck,	25	F	2	$ 8,158	10/9/2006	Home, Self-Care	--
26	373: Vaginal Delivery	MD6	373	3730	663.31: Oth Unspec Cord E	20	F	1	$ 4,596	10/9/2006	Home, Self-Care	--
27	373: Vaginal Delivery	MD6	373	3730	664.41: Unspec Perineal L	25	F	1	$ 7,449	10/9/2006	Home, Self-Care	Complication

Excel's PivotTable Tool

To summarize and analyze this data, you will want to learn how to use Excel's PivotTable tool. In past incarnations it was known as Crosstab (for cross tabulation). With Pivot Tables and the file above you could easily:

- Count the number of deliveries all doctors performed
- Count the number of times each doctor had a "Complication" during delivery.
- Sum or average the charges per delivery by doctor.
- Count the number of deliveries for each diagnosis.

Pivot tables are a great tool but the user interface is awkward. I have found that few people know how to use Excel's PivotTable function to analyze this kind of data. I have to believe that it's because the user interface isn't intuitive. But here's how you do it step by step.

Step 1: Your Data Must Have Column Headings!

As you can see in this data sheet, each column has a heading. *The Pivot Table will not run if there is a blank cell in any heading*. One of the first mistakes people make is inserting blank columns to make the file more readable, and then they wonder why the Pivot Table won't work. ***Avoid mistakes: No blanks in column headings!***

Step 2: Select the Data

You can select the data using your mouse or you can click on any cell in the data and the PivotTable will automatically select all of the data. Below I clicked on E6. Then when I click on Data-PivotTable to start the Pivot Table wizard, the wizard will automatically select all of the surrounding data – but I'm getting ahead of myself.

	A	B	C	D	E	F	G	H	I	J	K	L
1	DRG	Physi-cian ID #	DRG	APS DRG	Diagnosis	Age	Sex	LOS	Total Charge	Date	Discharge Status	Adverse Event(s)
2	373: Vaginal Delivery	MD11	373	3730	664.01: First-Degree Peri	26	F	2	$ 5,729	10/1/2006	Home, Self-Care	--
3	373: Vaginal Delivery	MD11	373	3730	645.11: Post Term Pregnan	18	F	2	$ 9,551	10/2/2006	Home, Self-Care	--
4	373: Vaginal Delivery	MD8	373	3730	663.31: Oth Unspec Cord E	37	F	1	$ 6,976	10/2/2006	Home, Self-Care	--
5	373: Vaginal Delivery	MD8	373	3730	650: Normal Delivery	19	F	1	$ 4,589	10/3/2006	Home, Self-Care	--
6	373: Vaginal Delivery	MD1	373	3730	650: Normal Delivery	28	F	2	$ 11,033	10/4/2006	Home, Self-Care	--
7	373: Vaginal Delivery	MD2	373	3730	663.31: Oth Unspec Cord E	27	F	1	$ 7,002	10/4/2006	Home, Self-Care	--
8	373: Vaginal Delivery	MD3	373	3730	646.81: Oth Spec Complica	24	F	2	$ 7,190	10/4/2006	Home, Self-Care	--
9	373: Vaginal Delivery	MD3	373	3730	645.11: Post Term Pregnan	21	F	2	$ 6,313	10/4/2006	Home, Self-Care	--
10	373: Vaginal Delivery	MD5	373	3730	650: Normal Delivery	19	F	1	$ 6,377	10/4/2006	Home, Self-Care	--
11	373: Vaginal Delivery	MD11	373	3730	656.61: Excessive Fetal G	22	F	1	$ 7,778	10/5/2006	Home, Self-Care	--
12	373: Vaginal Delivery	MD3	373	3730	664.01: First-Degree Peri	19	F	1	$ 6,753	10/5/2006	Home, Self-Care	--
13	373: Vaginal Delivery	MD6	373	3730	663.31: Oth Unspec Cord E	22	F	1	$ 8,369	10/5/2006	Home, Self-Care	Complication
14	373: Vaginal Delivery	MD6	373	3730	664.11: Second-Degree Per	24	F	2	$ 7,292	10/5/2006	Home, Self-Care	--
15	373: Vaginal Delivery	MD2	373	3730	664.01: First-Degree Peri	26	F	1	$ 6,464	10/6/2006	Home, Self-Care	Complication
16	373: Vaginal Delivery	MD3	373	3730	664.11: Second-Degree Per	20	F	1	$ 6,425	10/6/2006	Home, Self-Care	--

To initiate the PivotTable, click on Data and select Insert PivotTable:
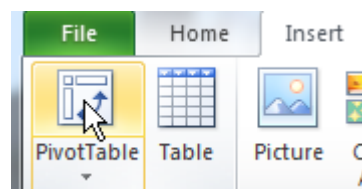
This will launch the PivotTable Wizard:
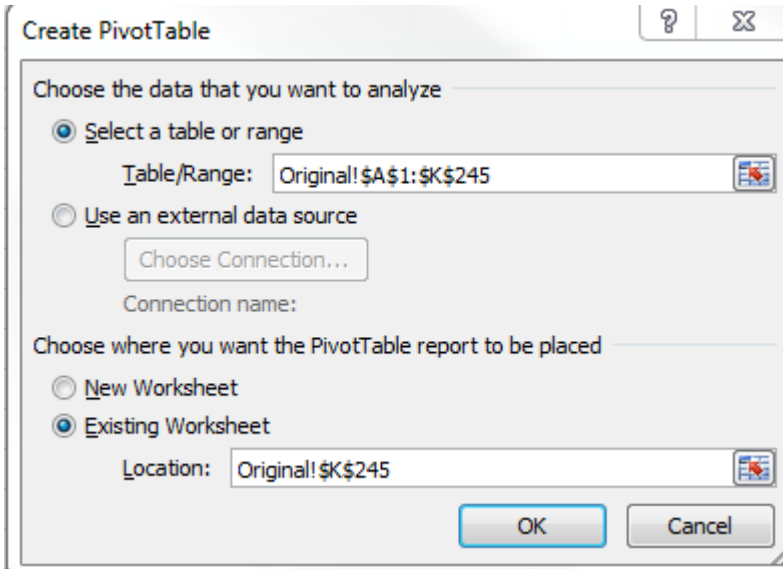
The most common defaults are selected: Microsoft Excel list or database and PivotTable. You could also pull in data from an external data source like Microsoft Access or you could summarize multiple ranges (i.e., more than one column contains words that you want to aggregate into one count).

I invariably just click Next > which leads us to Step 2 of the Wizard – selecting your data. If you've already clicked on a cell in the center of your data as I suggested, Excel will select everything around it; in this case A1:L240. I can choose to put the summaries in a new worksheet or the current worksheet.

Again, I usually just click OK! If you know what you're doing, you can usually use the Layout and Options to specify how you want the summaries to appear, but it's usually easier to work from the default table and field list:
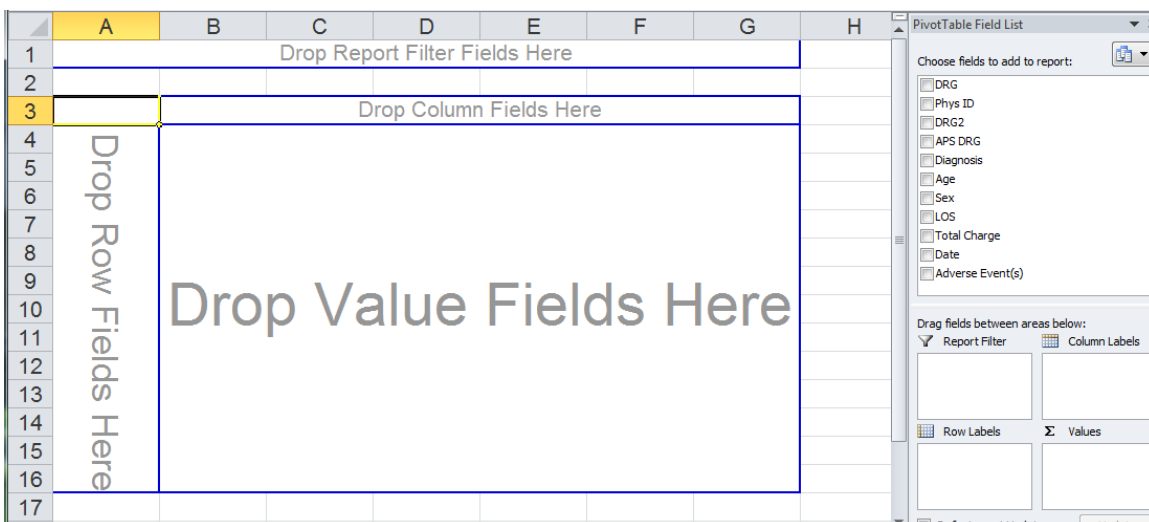
Step 3: Picking the Best Layout for Your Data

The PivotTable works on a "drag-and-drop" interface. Just click on a field and drag it into one of the four boxes:

- Page Fields for higher level summaries (e.g., facility or location names)
- Row Fields should have the most frequent heading (often dates)
- Column Fields should contain less frequently used headings.
- Data Items are where you drag and drop the words or numbers you want to count or summarize.

If I want to know how things add up by date, I usually drag the Date field into the Row Fields. Putting the dates down the left side prevents overflows and makes it easier to run a control chart with the QI Macros:

Then I might want to look at the results by Physician. So, I drag Physician into the Column Fields:

Then, maybe I want to summarize the Total Charges by physician by day; So I drag Total Charges into the Data Item Box which gives me totals by day and physician as well as some insight into how often the doctors work:

Date	MD1	MD10	MD11	MD2	MD3	MD5	MD6	MD8	Grand Total
10/01/16			5729						5729
10/02/16			9551					6976	16527
10/03/16								4589	4589
10/04/16	11033			7002	13503	6377			37915
10/05/16			7778		6753		15661		30192
10/06/16				6464	6425	7985		7299	28173
10/07/16								8344	8344
10/08/16				17368			19333		36701
10/09/16							20203		20203
10/10/16							7328		7328
10/11/16		6289						8061	14350
10/12/16					33551				33551
10/13/16		7855					27919		35774
10/14/16	4518		5382						9900
10/15/16				7391					7391
10/16/16				8260		8154			16414
10/17/16				5274		13204			18478
10/18/16						5627		5196	10823
10/19/16				8407			6736	4853	19996
10/20/16				6345			31487		37832
10/21/16			8949					6304	15253
10/22/16					6426			6508	12934
10/23/16				5886				14625	20511

PivotTable Field List

Choose fields to add to report:
- DRG
- ✓ Phys ID
- DRG2
- APS DRG
- Diagnosis
- Age
- Sex
- LOS
- ✓ Total Charge
- ✓ Date
- Adverse Event(s)

Drag fields between areas below:
- Report Filter
- Column Labels: Phys ID
- Row Labels: Date
- Σ Values: Sum of Total ...

Now let's say that I wanted to analyze the charges in terms of adverse events (pregnancy didn't go as planned). I'd just drag and drop Adverse Events into the Page Fields. The pivot table gives me a choice of viewing all charges or just the charges with the key word complication, Outlier, Readmission, etc.:

Date					MD2	MD3	MD5	MD6	MD8	Grand Total
Adverse Event(s)	(All)									
Sum of Total Charge										
10/1/2006										$ 5,729
10/2/2006									$ 6,976	$ 16,527
10/3/2006									$ 4,589	$ 4,589
10/4/2006					$ 7,002	$13,503	$ 6,377			$ 37,915
10/5/2006						$ 6,753		$15,661		$ 30,192
10/6/2006					$ 6,464	$ 6,425	$ 7,985		$ 7,299	$ 28,173
10/7/2006									$ 8,344	$ 8,344
10/8/2006					$17,368			$19,333		$ 36,701
10/9/2006								$20,203		$ 20,203
10/10/2006								$ 7,328		$ 7,328
10/11/2006		$ 6,289							$ 8,061	$ 14,350
10/12/2006						$33,551				$ 33,551
10/13/2006		$ 7,855						$27,919		$ 35,774
10/14/2006	$ 4,518			$ 5,382						$ 9,900

Dropdown list options:
- (All)
- Complication
- Complication, Outlier ($)
- Outlier ($)
- Readmission
- Readmission, Outlier ($)

If I select Outlier, I get some insight into the cost of complications. The costs are almost 50% higher than a normal delivery:

Adverse Event(s)	Outlier ($)				
Sum of Total Charge	Physi-cian ID #				
Date	MD1	MD11	MD2	MD3	Grand Total
10/25/2006	$ 13,805				$ 13,805
11/5/2006		$15,110			$ 15,110
11/12/2006			$13,082		$ 13,082
12/9/2006				$13,045	$ 13,045
Grand Total	$ 13,805	$15,110	$13,082	$13,045	$ 55,042

www.qimacros.com

I could also use Blank to evaluate all pregnancies with no complications:

Date	MD1	MD10	MD11	MD2	MD3	MD5	MD6	MD8	Grand Total
		Physi-c ▾							
10/1/2006			$ 5,729						$ 5,729
10/2/2006			$ 9,551					$ 6,976	$ 16,527
10/3/2006								$ 4,589	$ 4,589
10/4/2006	$ 11,033			$ 7,002	$ 13,503	$ 6,377			$ 37,915
10/5/2006			$ 7,778		$ 6,753		$ 7,292		$ 21,823
10/6/2006					$ 6,425	$ 7,985		$ 7,299	$ 21,709
10/7/2006								$ 8,344	$ 8,344
10/8/2006				$ 17,368			$ 19,333		$ 36,701
10/9/2006							$ 12,754		$ 12,754
10/10/2006							$ 7,328		$ 7,328
10/11/2006		$ 6,289						$ 8,061	$ 14,350
10/12/2006					$ 24,393				$ 24,393
10/13/2006							$ 19,826		$ 19,826
10/14/2006			$ 5,382						$ 5,382

Row 3: Sum of Total Charge Physi-c ▾

If you double click on "Sum of Total Charge" and change it to Average, you get the average:

Date	MD1	MD10	MD11	MD2	MD3	MD5	MD6	MD8	Grand Total
10/1/2006			$ 5,729						$ 5,729
10/2/2006			$ 9,551					$ 6,976	$ 8,264
10/3/2006								$ 4,589	$ 4,589
10/4/2006	$ 11,033			$ 7,002	$ 6,752	$ 6,377			$ 7,583
10/5/2006			$ 7,778		$ 6,753		$ 7,292		$ 7,274
10/6/2006					$ 6,425	$ 7,985		$ 7,299	$ 7,236
10/7/2006								$ 8,344	$ 8,344
10/8/2006				$ 8,684			$ 6,444		$ 7,340
10/9/2006							$ 6,377		$ 6,377
10/10/2006							$ 7,328		$ 7,328
10/11/2006		$ 6,289						$ 8,061	$ 7,175
10/12/2006					$ 8,131				$ 8,131
10/13/2006							$ 6,609		$ 6,609
10/14/2006			$ 5,382						$ 5,382

Row 3: Average of Total Charge Physi-c ▾

I could also change it back and group all of the charges into monthly charges. How to: Just Right click on any date and select Group and Show Detail/Group/Months:

Row 3: Average of Total Charge Physi-cian ID # ▾
Row 4: Date ▾ MD1   MD10

Context menu:
- Format Cells...
- PivotChart
- Wizard...
- Refresh Data
- Hide
- Select ▸
- Group and Show Detail ▸
  - Hide Detail
  - Show Detail
  - Group...
  - Ungroup...
- Order ▸
- Field Settings...
- Table Options...

Grouping dialog:
Auto
☑ Starting at: 10/1/2006
☑ Ending at: 1/1/2007
By
Seconds
Minutes
Hours
Days
Months
Quarters
Number of days: 1
[OK] [Cancel]

The Pivot Table will group the months.

	A	B	C	D	E	F	G	H	I	J
1	Adverse Event(s)	--								
2										
3	Sum of Total Charge	Physi-cia								
4	Date	MD1	MD10	MD11	MD2	MD3	MD5	MD6	MD8	Grand Total
5	Oct	$ 24,828	$ 31,100	$ 52,869	$ 45,008	$ 100,272	$ 34,655	$ 150,798	$ 86,293	$ 525,823
6	Nov	$ 6,990	$ 64,650	$ 37,244	$ 27,968	$ 70,086	$ 34,651	$ 111,462	$ 89,923	$ 442,974
7	Dec	$ 44,991	$ 48,929	$ 82,657	$ 24,290	$ 94,903	$ 53,542	$ 77,620	$ 89,109	$ 516,041
8	Grand Total	$ 76,809	$ 144,679	$ 172,770	$ 97,266	$ 265,261	$ 122,848	$ 339,880	$ 265,325	$1,484,838

Warning! Excel Bonehead Behavior: Grouping dates will not work if there is even a single blank or text cell where there should be a date.

Now I can select B4:I4, hold down the control key to select B8:I8 and run a Pareto chart:

Total Charges by Physician
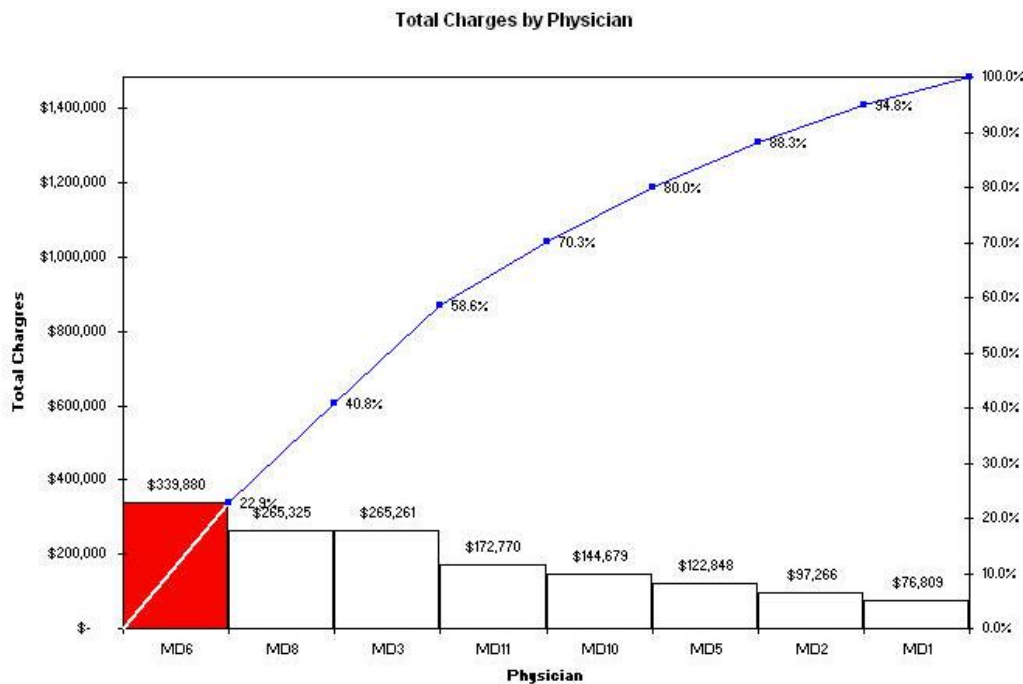
Changing the Focus

What if I wanted to change this table to count adverse events by Physician and by age of patient? *Hint: It's easy with drag and drop.*

1. Just click on date and pull it out of the table.

2. Click on Sum of total Charge and pull it out of the table.

3. Drag Adverse events down into Data Fields.

4. Then click on the Age field and drag it into the Row Fields:

	A	B	C	D	E	F	G	H	I
1	Adverse Event(s)	Complication ▼							
2									
3	Count of Adverse Event(s)	Physi-cian ID # ▼							
4	Age ▼	MD1	MD10	MD2	MD3	MD5	MD6	MD8	Grand Total
5	18				1		1		2
6	19					1	2		3
7	20				1		1		2
8	21	1					1	1	3
9	22						3		3
10	23				1				1
11	24	1			1				2
12	25						1		1
13	26			1					1
14	27	1							1
15	33		1						1
16	37				1				1
17	38			1					1
18	Grand Total	3	1	2	5	1	9	1	22

Are there more younger women with complications or are there just more younger women giving birth? Does one physician have more complications than the others? We could run a Pareto chart to show complications by physician (MD6 is 40% of total complications, almost twice as high as his or her peers):

Complications by Physician
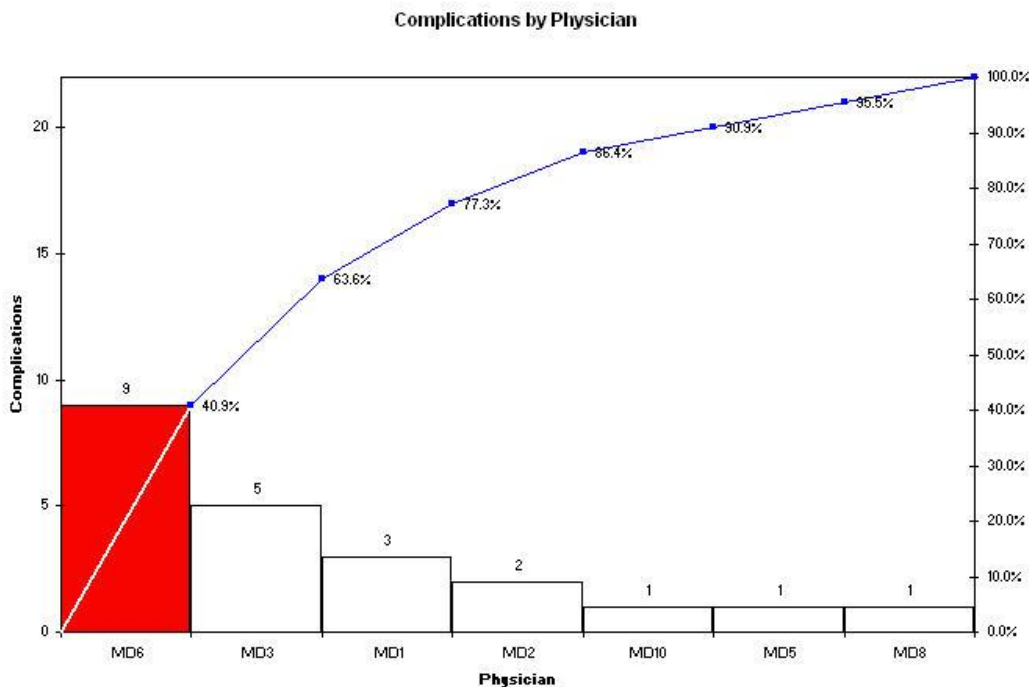
What Else?

Maybe I'd like to evaluate average length of stay (LOS) for all deliveries. Just pull the adverse events out of the table and drop in LOS (change it to an average). Not much going on here. The youngest and the oldest had a slightly longer length of stay:

Adverse Event(s)	(All)								
Average of LOS	Physi-cian ID #								
Age	MD1	MD10	MD11	MD2	MD3	MD5	MD6	MD8	Grand Total
14		2.00		2.00			2.00		2.00
15							1.50		1.50
17							2.00	1.00	1.50
18		1.00	2.00		1.75		1.50		1.54
19		2.00	1.33		1.33	1.33	2.00	1.60	1.58
20		2.00			1.40	1.00	1.29	1.50	1.41
21	1.50	1.33	1.00	1.00	1.67	1.00	1.50	1.57	1.45
22	1.00	1.50	1.00	1.00	1.71	1.00	1.71	2.00	1.58
23			1.33	3.00	2.00	1.67	1.00	1.67	1.73
24	1.00		1.50		1.50	2.00	2.50		1.75
25	1.00	1.00			2.00	1.33	1.50	2.00	1.53
26		1.50	1.33	1.50	1.75		1.50	1.25	1.47
27	2.00	2.00	1.00	1.00	1.00		1.00	1.40	1.27
28	2.00	1.50	1.25	1.00		2.00	1.25		1.43
29	1.50		2.00		2.00		1.00	1.50	1.57
30		2.00	1.00	2.00	1.00	1.00		1.00	1.33
31	2.00	1.00		1.00	2.00			2.00	1.60
32	2.00			1.33	2.00		2.00	2.00	1.78
33		1.00		1.00	1.00		1.00		1.00
34	1.50		1.00				2.00		1.60
35	1.00				1.00		1.50	1.00	1.20
36		1.00						1.00	1.00
37					2.00			1.00	1.33
38				1.00		2.00			1.50
40	2.00								2.00
Grand Total	1.53	1.43	1.36	1.44	1.65	1.44	1.55	1.49	1.51

Get the Idea?

There's a wealth of information hiding in these dense flat files of words and numbers. Start using Excel's Pivot Table function to slice and dice your files (no matter how large). Then use the QI Macros to graph the results. You'll find it easy to use the 4-50 Rule and start making breakthrough improvements.

QI Macros Improvement Project Wizard Will Do All of This Analysis for You

The Improvement Project Wizard will take your data and create improvement projects. Just select two headings (e.g., dates and charges) and click on the QI Macros Improvement Project Wizard.

File	Home	Insert	Page Layout	Formulas	Data	Review	View	Developer	QI Macros 2021		
Stat Wizard	Lean Tools ▾	Improvement Tools ▾		Improvement Project Wizard	Box, Dot & Scatter Plot ▾		Control Chart Wizard		Pareto Chart		Chart
Statistical Tools ▾	Diagrams ▾	DOE GageR&R FMEA ▾		Data & Text Mining	Histograms & Capability ▾		Control Charts (SPC) ▾		Fishbone ▾		Run C
Stat Templates ▾	Calculators ▾	Planning & PM Tools ▾		Restacking ▾	Capability Templates ▾		Control Chart Templates ▾		Chart Templates ▾		Excel Cha
Statistical Tools		Templates for Lean Six Sigma			Data Mining	Capability Charts		Control Charts		Improve Charts		Other C

J1 ▾ *fx* Date

	A	B	C	D	E	F	G	H	I	J	K
1	DRG	Phys ID	DRG	APS DRG	Diagnosis	Age	Sex	LOS	Total Charge	Date	Adverse Event(s)
2	373: Vaginal Delivery	MD1	373	3730	650: Normal Delivery	28	F	2	$ 11,033	10/04/16	
3	373: Vaginal Delivery	MD1	373	3730	661.21: Oth Unspec Uterin	24	F	1	$ 4,518	10/14/16	Complication
4	373: Vaginal Delivery	MD1	373	3730	661.21: Oth Unspec Uterin	35	F	1	$ 5,188	10/24/16	

The Wizard would have found four possible projects in this data:

	A	B	C
1	QI Macros Analysis	Charts and Diagrams	Problem Statements
2			
3	Control Chart	Total Charge by Date	
4			
5	Project 1		
6	Pareto Chart 80-20 Rule	Total Charge by Diagnosis	
7			
8	Project 2		
9	Pareto Chart 80-20 Rule	Total Charge by LOS	
10			
11	Project 3		
12	Pareto Chart 80-20 Rule	Total Charge by Phys ID	
13			
14	Project 4		
15	Pareto Chart 80-20 Rule	Total Charge by Adverse Event(s)	
16			
17	Histogram	Original no breaks Total Charge Total Charge	
18			

Need More Help or Don't Have the Time?

- We can look at your data and quickly tell if there's an improvement project lurking in the rows and columns of your Excel worksheet. Here's How It Works:
- Just send us your Excel workbook with a description of the issue you're trying to resolve.
- We will look at the data you've sent. If we don't think there's an improvement project in your data, we'll tell you, no charge.
- If there are one or more projects in your data, we'll develop the improvement project for you.

Click here for more details

COUNTIF

If you have comment fields filled with text, use Excel's COUNTIF function. COUNTIF will count the number of times a text expression occurs in a range of cells. The following data shows comments about insurance billing problems. Maybe we'd like to know how many are related to Medicare or how many are related to *overlapping dates of stay (DOS)*. The COUNTIF Function can find and count any of these:

A1		=	'DC REJECT - OVERLAPPING DOS. - MDCR HAS TAKEN BACKORG PYMT FOR DOS 4.19.02-4.30.02 DUE TO INPT OVERLAP. CLD MDCR TT RYAN, THIS TAKE BACK WAS AN ERROR...		
1	DC REJECT - OVERLAPPING DOS. - **MDCR** HAS TAKEN BACKORG PYMT FOR DOS 4.19.02-4.30.02 DUE TO INPT OVERLAP. CLD **MDCR** TT RYAN, THIS TAKE BACK WAS AN ERROR!	Keyword	COUNTIF	
2	DC REJECT - OVERLAPPING DOS. CLAIM DENIED FOR IP/OP PROXIMITY. NO ANSWER WHEN CALLED. LETTER WASSENT REQUESTING INFO ON IP STAY. IT WAS NOT AT ONEOF OUR FACILITIES.	m*d*c*r	74 =COUNTIF(A:A,"=*"&B2&"*")
3	DC REJECT - OVERLAPPING DOS. **MCR** DENIED 092101 FOROVERLAPPING CHRGS. REFER TO MAX FOR CHRG TRANSFERTO ACCT (001317832-1243).	medicare	5 =COUNTIF(A:A,"=*"&B3&"*")
4	DC REJECT - OVERLAPPING DOS. **MCR** DENIED 061002 FORINPT/OUTPT OVERLAP. REFER TO MAX FOR CHRGE TRANSFER TO ACCT 001640742-0001.	mcr	25
5	DC REJECT - OVERLAPPING DOS. PER **MCARES** DENIAL, WILL TRANSFER THESE CHARGES TO INPT EPISODE 1109. WILL REFER TO INPT F/U. /	mdcr	18
6	DC REJECT - OVERLAPPING DOS. REQ FROM MAX A BALANCE TRANSFER FOR THIS STAY TO IP STAY 0006. **MEDCR** DENIED FOR IN/OUT PROXIMITY.	medcr	7
7	DC REJECT - OVERLAPPING DOS. INPT #1029 HAS BEEN RESTORED FROM THE ARCHIVED FILES. REQUESTED A CHRG TRANSFER IN THE AMNT OF $1439.82.	mcare	6
8	DC REJECT - OVERLAPPING DOS. - **MDCR** HAS DENIED DOS 9.10.02-9.23.02 DUE TO AN INPT CLAIM DOS 9.16.02-9.17.02..	overlap*dos	175
9	DC REJECT - OVERLAPPING DOS./TRANSFERED CHARGES TOEPI 0005 AS **MEDICARE** PAID THAT EPI AND DENIED 0004/		
10	DC REJECT - OVERLAPPING DOS. I HAVE REQ'D THIS CHGS OF 123.20 TO BE TRANSFERRED 123.20 TO VISIT #0002.		

As you can see in this example, billing analysts spell Medicare in numerous different ways. Using COUNTIF, we can count the number of occurrences of each abbreviation. COUNTIF takes the form:

=COUNTIF(range,criteria) = COUNTIF(A:A,"=*"&B2&"*")

In this case, I'm using column A (A:A) as the range. And I'm using a formula to create the criteria based on values I input into cells in column B. I find this is easier than embedding the criteria in the COUNTIF, because I'm always refining the keywords and I don't want to change the formulas, just the text.

In this example, COUNTIF will look for every cell that begins with any number of characters ("=*"), followed by whatever is in the keyword cell (&B2&), followed by any number of characters ("*"). The asterisk is a *wildcard character that matches zero or more characters).*

Wildcard characters you can use to find text or numbers

To find text or numbers that have some characters or digits in common, use a wildcard character. A wildcard character represents one or more unspecified characters.

Use	To find
? (question mark)	Any **single** character in the same position as the question mark. For example, "sm?th" finds "smith" and "smyth"
* (asterisk)	Any number of characters in the same position as the asterisk. For example, "*east" finds "Northeast" and "Southeast"
~ (tilde) followed by ?, *, or ~	A question mark, asterisk, or tilde For example, What~? finds "What?"

I could also use the wild card characters to match any occurrence of M*D*C*R which will give me a close approximation to the total number of Medicare errors.

I might also want to know how many were *denied* and how many were *transferred*. It's easy to add keywords and counts using COUNTIF and the keyword formula.

7	$1439.82.	mcare	6
8	DC REJECT - OVERLAPPING DOS. - **MDCR** HAS DENIED DOS 9.10.02-9.23.02 DUE TO AN INPT CLAIM DOS 9.16.02-9.17.02..	overlap*dos	175
9	DC REJECT - OVERLAPPING DOS./TRANSFERRED CHARGES TO EPI 0005 AS **MEDICARE** PAID THAT EPI AND DENIED 0004/	denied	62
10	DC REJECT - OVERLAPPING DOS. I HAVE REQ'D THIS CHGS OF 123.20 TO BE TRANSFERRED 123.20 TO VISIT #0002.	transfer	85

These counts can then be graphed as bar charts (to show comparison in size) or Pareto charts using the QI Macros. These can become part of your dashboard as well.

As you can see, this makes comment field analysis much simpler than most people can imagine. Where pivot tables work on fields that have fixed values, COUNTIF works on fields with variable input.

The QI Macros Word Count Wizard will do this for you automatically. Just select the data and click on the Word Count Wizard:
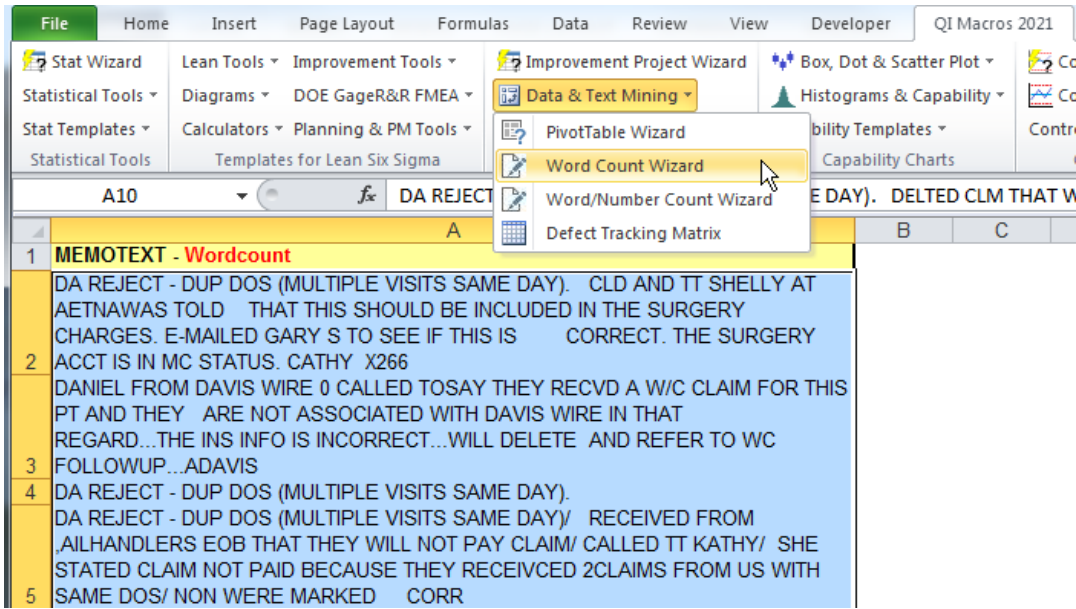
The Wizard will parse this text into one- and two-word phrases and summarize them. "Dup dos," "dos multiple" and "multiple visits" all point to the same problem:

	A	B	C	D	E
1	Count of Word			Count of Two-Word Phrases	
2	Word	Total		Two-Word Phrases	Total
3	dup	11		da reject	8
4	dos	9		reject dup	8
5	reject	8		dup dos	8
6	da	8		multiple visits	6
7	same	7		same day	6
8	multiple	6		dos multiple	6
9	day	6		visits same	6