

# 10 Tips for Better Excel Results

Peter Gascoyne

[PeterG@MakingApples.com](mailto:PeterG@MakingApples.com)

**This short presentation will provide 10 pointers to help improve your work & results using MS Excel.**

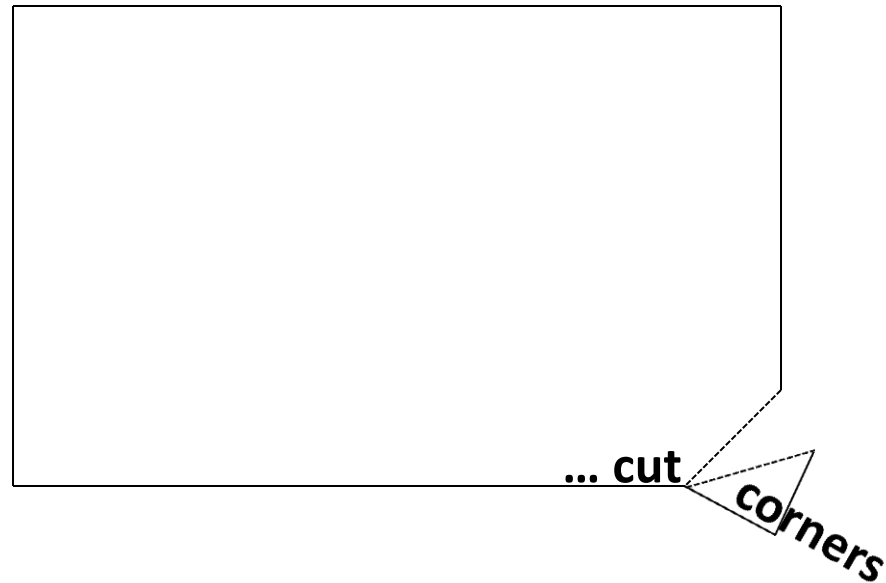
- 1. The most common mistake analysts make.**
- 2. Think about basic file structure.**
- 3. Set up table rows & columns in even increments.**
- 4. Don't imbed values in formulas.**
- 5. Use conditional formatting to ease reading.**
- 6. Apply the "2-4 Digit Rule".**
- 7. Array formulas are a very powerful tool.**
- 8. Find & Replace can speed up work.**
- 9. Charts with 2 axes should be proportional.**
- 10. Use multiple pages (& charts) to tell story well.**

Hi, my name is Peter Gascoyne. I've been a financial & economic consultant for almost 30 years. Over this time, as you might imagine, I've developed a lot of spreadsheets and models to run the analysis on all the various projects I've

been involved. In this short presentation, I am going to draw upon that experience, and walk through 10 tips that I think will help improve the analytical work you perform using Microsoft Excel. My focus here will be on some basic ideas

on how to approach and perform your analysis. The emphasis will be on the ideas, not the precise step-by-step procedures & formulas.

# 1. The most common mistake analysts make is to...



Perhaps the biggest mistake I see analysts make is to cut corners, to attempt the shortcuts that will get the project completed “more quickly”. Time & time again these shortcuts lead to more work, more effort, more time expended to get the project done. How often have you rushed a job, only to find yourself re-doing the work because it wasn’t right the 1<sup>st</sup> time, or errors cause management to distrust your product ? How often has the product been wrong because of the rush?

We live in stressful times, when all too frequently forces to bear are pushing for results sooner, rather than later. It can be habit-forming; one rushes because one has always rushed. And all too often we fall prey to a tendency to focus on the more enjoyable & satisfying “analysis”, and are reluctant to conduct the more tedious, yet vital data prep & data adjusting. Data is noisy, and invariably its noise masks its message. To clarify the message, to cull out the story it has to tell, frequently requires significant time & effort adjusting & preparing the data.

Accounting errors may have been made – they need to be found, and corrected. Sales may jump due to the latest ad campaign – but due consideration & adjustments are also required if the weather was particularly obliging, and the sales staff motivated by larger than usual bonus. Results don’t look right – if they don’t, they need to be scrutinized, even if the deadline is near, or a tired colleague pronounces them fine.

Save time; take your time.

## 2. Think about the basic structure of your files: have separate tabs with very distinct functions.

### Example:

1. **Input**      **Data, assumptions, notes, etc.**
2. **Calc**      **Key tables & calculations**
3. **Output**      **Summary tables, charts, etc.**

Give some thought to the basic structure you employ with *all* your files. Above I've outlined an approach I've found has worked well. I try to have most files contain three (or more) tabs. The first is an "Input" tab, where all the data past, present, & future is brought in. A second tab is the "Calc" tab, which is the workhorse of the file, where almost all the tables & calculations are

performed. Generally speaking, once established, the "Calc" tab can be left alone – the effort is on bringing new data into the "Input" tab, & observing its impact in the third tab, the "Output" tab. The "Output" tab would have the summary tables that get copied into PowerPoints, and the charts that capture the story.

Obviously, one can play with this, having separate tabs for summary tables and summary charts, separate tabs for data and for notes, and so on. My point is to just think about how you want to structure your files as a rule – is there a way you can think of that will make them easier to work with in future?

### 3. Set up tables in easy to work with increments.

Break tables into even-numbered rows: 10, 20, 100, 1,000 ...

	A	B	C	D	E
1	<b>Product Sales (\$ Millions)</b>				
2					
9				<b>2016</b>	
10	<b>Sales</b>				
11		U.S.	1,142		
12		Americas	751		
13		Europe	1,375		
14		Africa & MidEast	629		
15		Asia	1,743		
16		<b>Total</b>	<b>5,640</b>		
17					
20	<b>Plan</b>				
21		U.S.	1,273		
22		Americas	725		
23		Europe	1,312		
24		Africa & MidEast	606		
25		Asia	1,786		
26		<b>Total</b>	<b>5,702</b>		
27					
30	<b>Variance</b>				
31		U.S.	(131)		
32		Americas	26		
33		Europe	63		
34		Africa & MidEast	23		
35		Asia	(43)		
36		<b>Total</b>	<b>(62)</b>		

This suggestion may seem silly, but I've really found it helpful over the years to structure worksheets into clear, easy-to-work-with increments. Once you establish a pattern it becomes easier to remember & locate things. In the example here, Europe is always found in a row ending in "3". To reduce empty space &

improve the visual, blank rows are hidden. Of course, depending on what you're working with, different increments will be appropriate. Increments of 10, 100, 1,000 rows work best, but reducing hidden row counts may prompt using 20, 40, 50, 200, etc. as the pattern.

I might add I've also found it handy using the top 10 rows for headings, assumptions, etc. It provides plenty of space, it can be easily applied to almost any worksheet (instead of constantly starting tables at different-numbered rows), and it works well with the row-numbering structure to begin at Row 11.

# Going across a spreadsheet, it can be helpful to break out tables into increments of 26 columns.

	A	B	C	D	E	F	W	X	AB	AC	AD	AE	AF	AW	BB	BC	BD	BE	BF	BW	
1	<b>Product Sales (\$ Millions)</b>																				
2																					
7			<b>2016</b>							<b>2017</b>							<b>2018</b>				
9			Jan-15	Feb-15	Mar-15	Q1	<b>2016</b>		Jan-16	Feb-16	Mar-16	Q1	<b>2017</b>		Jan-17	Feb-17	Mar-17	Q1	<b>YTD '18</b>		
10	<b>Sales</b>																				
11	U.S.		69	80	103	251	1,142		65	77	101	243	1,184		78	86	110	274	554		
12	Americas		45	53	68	165	751		43	51	67	161	786		43	50	69	162	330		
13	Europe		83	96	124	303	1,375		78	92	121	292	1,423		76	90	114	280	566		
14	Africa & MidEast		38	44	57	138	629		36	43	56	135	658		38	40	52	130	266		
15	Asia		105	122	157	383	1,743		104	123	161	388	1,893		111	120	153	384	774		
16	<b>Total</b>		<b>338</b>	<b>395</b>	<b>508</b>	<b>1,241</b>	<b>5,640</b>		<b>327</b>	<b>386</b>	<b>505</b>	<b>1,219</b>	<b>5,944</b>		<b>346</b>	<b>386</b>	<b>498</b>	<b>1,230</b>	<b>2,490</b>		
17																					
20	<b>Plan</b>																				
21	U.S.		76	89	115	280	1,273		66	78	103	247	1,206		74	87	113	274	535		
22	Americas		44	51	65	160	725		41	49	64	154	753		42	49	63	154	300		
23	Europe		79	92	118	289	1,312		77	92	120	289	1,409		77	90	116	283	552		
24	Africa & MidEast		36	42	55	133	606		35	42	55	132	642		34	40	53	127	249		
25	Asia		107	125	161	393	1,786		103	122	160	386	1,881		105	123	159	387	756		
26	<b>Total</b>		<b>342</b>	<b>399</b>	<b>513</b>	<b>1,254</b>	<b>5,702</b>		<b>324</b>	<b>383</b>	<b>501</b>	<b>1,208</b>	<b>5,891</b>		<b>332</b>	<b>389</b>	<b>503</b>	<b>1,224</b>	<b>2,392</b>		
27																					
30	<b>Variance</b>																				
31	U.S.		(8)	(9)	(12)	(29)	(131)		(1)	(1)	(2)	(5)	(22)		4	(1)	(3)	(0)	19		
32	Americas		2	2	2	6	26		2	2	3	7	33		2	1	6	8	30		
33	Europe		4	4	6	14	63		1	1	1	3	14		(1)	0	(2)	(3)	14		
34	Africa & MidEast		1	2	2	5	23		1	1	1	3	16		4	(0)	(1)	3	17		
35	Asia		(3)	(3)	(4)	(9)	(43)		1	1	1	2	12		6	(3)	(6)	(3)	18		
36	<b>Total</b>		<b>(4)</b>	<b>(4)</b>	<b>(6)</b>	<b>(14)</b>	<b>(62)</b>		<b>3</b>	<b>3</b>	<b>5</b>	<b>11</b>	<b>53</b>		<b>14</b>	<b>(3)</b>	<b>(5)</b>	<b>6</b>	<b>98</b>		

Just as with rows, breaking out columns of info into increments of 26 columns can be very helpful. In this example, where quarter totals are shown, and where April thru December values are not visible due to display space for this page, you

can see that January values are always in the "C" columns (C, AC, BC, etc.). Similarly, annual totals are found in the "W" columns. Why bother with all this? If you're on another tab

in the file, or you're reviewing formulas for errors, it can be extremely useful to rely on certain info always being located in a given column. You know totals come from a "W" column; if they aren't, you've found a mistake.

# 4. Do not imbed values in formulas.



	A	B	C	D	E
1	<b>Formula Using Values</b>				
2					
10		<u>Period 1</u>	<u>Period 2</u>	<u>Period 3</u>	<u>Period 4</u>
11	Overhead	100.0	103.0	106.1	109.3

Formula: = B11 \* 1.03



	A	B	C	D	E
1	<b>Formula Using Cell References</b>				
2					
8	Inflation:	3.0%			
9					
10		<u>Period 1</u>	<u>Period 2</u>	<u>Period 3</u>	<u>Period 4</u>
11	Overhead	100.0	103.0	106.1	109.3

Formula: = B11 \* (1+\$B\$8)

As much as possible, I would recommend resisting the temptation to use values in your formulas. For example, in the top figure, Period 2 is applying 3% inflation to overhead expenses by multiplying Period 1's amount by the value "1.03". There are a number of problems with imbedding values in formulas:

- It can be unclear where the value is coming from or what it represents, a problem that

expands with time.

- It can be harder to trace errors if your spreadsheet isn't working as it should.
- If you later want to change the value, it can be cumbersome to make all the revisions, & you risk not applying all the changes that need to be made.
- It can be difficult to tell in a printout how all the numbers are derived.

Instead of using values, have formulas reference other cells. That way you can better tell what is going on in the spreadsheet. You can use formatting to highlight key variables, like inflation. And you can easily change the input.

Nothing is hidden.

# 5. Conditional formatting can really help make a spreadsheet easier to read, and lets you highlight what you want.

**Call Center Performance** (showing Conditional Formatting)

	Jan	Feb	Mar	Apr	YTD
Count of Bad Evaluations					
Alex	5	8	3	1	17
Brenda	3	2	0	1	6
Carol	0	3	1	2	6
Doug	0	0	2	0	2
% Change in Call Volume Handled (Year-Over-Year)					
Alex	-2%	2%	-2%	0%	0%
Brenda	1%	-3%	0%	4%	1%
Carol	3%	4%	-1%	1%	3%
Doug	4%	-1%	2%	2%	3%

**Grey "0"s** (points to cells with 0)

**Negative Values in Red** (points to negative percentage cells)

**Format Based on Formula** (points to YTD column with conditional formatting rules: > 12 (red), < 4 (green))

Excel's conditional formatting can really help make spreadsheets easier to read, and allow you to automatically highlight information meriting attention. This example demonstrates three types of conditional formatting I've found helpful.

- Some spreadsheets contain lots of zeroes – by formatting them in grey they slip more into the background, where they belong.
- Showing negative values in red, and/or in brackets, make negative numbers pop out more quickly. This is especially helpful for percentages; yes, custom formats can also display them in red, but they can be more cumbersome to work with.
- Column H in this example shows formulas applying conditional formatting. The formatting was set so that whenever anyone's total is above the "12" input in Cell H7, the cell fills red, while if their total is below the "4" in Cell H8, the cell fills green. Changing the high value in Cell H7 from "12" to "20" would instantly cause Alex's total of 17 to lose the red highlight.



# 6. The “2-4 Digit Rule” expresses values so they are easier to read, easier to work with, and just make sense.



	<u>2014</u>	<u>2015</u>	<u>2016</u>	<u>2017</u> (Estimate)	<u>2018</u> (Forecast)
<b>Expenses</b>					
Labor	\$11,472,375	\$11,736,240	\$11,982,701	\$12,282,268	\$12,564,760
Energy	1,854,617	1,897,273	1,937,116	1,985,544	2,031,211
Maintenance	942,391	964,066	984,311	1,008,919	1,032,124
Other	2,614,521	2,674,655	2,730,823	2,799,093	2,863,472
<b>Total</b>	<b>\$16,883,904</b>	<b>\$17,272,234</b>	<b>\$17,634,951</b>	<b>\$18,075,824</b>	<b>\$18,491,568</b>



	<u>2014</u>	<u>2015</u>	<u>2016</u>	<u>2017</u> (Est)	<u>2018</u> (Forecast)
<b>Expenses (\$ MM)</b>					
Labor	\$11.5	\$11.7	\$12.0	\$12.3	\$12.6
Energy	1.9	1.9	1.9	2.0	2.0
Maintenance	0.9	1.0	1.0	1.0	1.0
Other	2.6	2.7	2.7	2.8	2.9
<b>Total</b>	<b>\$16.9</b>	<b>\$17.3</b>	<b>\$17.6</b>	<b>\$18.1</b>	<b>\$18.5</b>



	<u>2014</u>	<u>2015</u>	<u>2016</u>	<u>2017</u> (Est)	<u>2018</u> (Forecast)
<b>Expenses (\$ MM)</b>					
Labor	\$11.472	\$11.736	\$11.983	\$12.282	\$12.565
Energy	1.855	1.897	1.937	1.986	2.031
Maintenance	0.942	0.964	0.984	1.009	1.032
Other	2.615	2.675	2.731	2.799	2.863
<b>Total</b>	<b>\$16.884</b>	<b>\$17.272</b>	<b>\$17.635</b>	<b>\$18.076</b>	<b>\$18.492</b>

The “2-4 Digit Rule” argues for values being expressed with somewhere between 2 to 4 digits. Granted, for accounting purposes, you may need to track figures to a very precise level. Fine. But for presentation purposes, showing 2-4 digits, as a rule, is plenty. Why?

- The more digits, the less the reader actually sees. It’s very hard mentally to retain any more than 2-4 digits. Less is more.
- Beyond 3-4 digits, the numbers shown are

meaningless. If they don’t mean anything, why show them?

- Aside from the visual aesthetics, often it’s just inappropriate. For example, forecasts are estimates by nature. To show a lot of digits suggests a precision that looks unprofessional. It makes sense to say you think 2018’s labor costs will be about \$12.6 million; it does not make sense to say you think they will be \$12,564,760, or “about” that amount.

Certainly there can be occasions when it is appropriate to “break” the rule. The 3<sup>rd</sup> table here goes to 5 digits for some of the values, so that one can better see how costs are increasing from one year to the next.

Ultimately, it is a judgement call. But consider your audience, and assist them by displaying as few numbers as reasonably possible.

# 7. Excel's "Find & Replace" function can be a very powerful tool.

	A	B	C	D	E	F
1		<b>Product Sales (\$ Millions)</b>				
2						
9			<u>2016</u>	<u>2017</u>	<u>2018</u>	
10		<b>Sales</b>				
11		U.S.	1,142	1,184	554	
12		Americas	751	786	330	
13		Europe	1,375	1,423	566	
14		Africa & MidEast	629	658	266	
15		Asia	1,743	1,893	774	
16		Total	5,640	5,944	2,490	
17						
20		<b>Plan</b>				
21		U.S.	1,273	1,206	535	
22		Americas	725	753	300	
23		Europe	1,312	1,409	552	
24		Africa & MidEast	606	642	249	
25		Asia	1,786	1,881	756	
26		Total	5,702	5,891	2,392	
27						
30		<b>Variance</b>				
31		U.S.	(131)	(22)	19	
32		Americas	26	33	30	
33		Europe	63	14	14	
34		Africa & MidEast	23	16	17	
35		Asia	(43)	12	18	
36		Total	(62)	53	98	
37						
38						
39						
40		<b>U.S.</b>				
41		Sales	1,142	1,184	554	
42		Plan	1,273	1,206	535	
43		Variance	(131)	(22)	19	

=C\$11    =D\$11    =E\$11

**Copy Row 41 to Row 42:43. Then...**

**For "Plan" row, Find "\$1" and Replace with "\$2".**

**For "Variance" row, Find "\$1" and Replace with "\$3".**

I'm not going to get into a lot of Excel functions & formulas here; there are numerous useful resources available; (I particularly recommend Chandoo.org). But if you're fairly new to the world of analysis & Excel, there are two features I recommend you become familiar with.

The first is the Find & Replace function, quickly obtainable by entering "Ctrl+F" (or

Home:Editing:Find& Select). The function allows you to find a given entry and replace it with something else, for whatever range you've highlighted. In the simple example here, drawn from Tip #3, I've put formulas with "\$" signs to fix the row number, to pick up U.S. Sales. I then copy that formula down to Rows 42 & 43. Then, in Row 42, using Ctrl+F, I replace "\$1" with "\$2" to pick up "Plan" amounts, & in Row 43, replace

"\$1" with "\$3" to pick up "Variance" amounts.

Obviously, in this simple example the Find & Replace function seems cumbersome, but as your worksheet grows in size & complexity, it can prove more & more handy. You will find the judicious use of "\$" signs in your formulas will expand the use of this powerful Excel function.

# 8. Array formulas allow you to pull & calculate data of almost any selected criteria.

	A	B	C	D	E	F	G	H
1	<b>Using Array Formulas</b>							
2								
9		Sales by Period						
10		Period 1	Period 2	Period 3	Period 4		Totals	
11	James	26	6	49	39		120	
12	Bill	31	35	37	34		137	
13	Mary	22	7	10	27		66	
14	Mary	30	36	12	14		92	
15	James	12	25	37	46		120	
16	James	21	28	6	43		98	
17	Bill	14	4	6	28		52	
18	Bill	22	27	39	35		123	
19	Bill	40	19	22	16		97	
20	Mary	44	15	42	26		127	
21	James	49	41	30	37		157	
22								
23								
24	Totals	<	100					
25	James		21	28	6	43		98
26	Bill		54	23	28	44		149
27	Mary		52	43	22	41		158
28								
29	Totals	=>	100					
30	James		87	72	116	122		397
31	Bill		53	62	76	69		260
32	Mary		44	15	42	26		127

Example picks up sales above for "James", when his Total was less than the highlighted amount of "100".

**Formula in Cell C25:**  
`{=SUM((($B25=$B$11:$B$21) * ($H$11:$H$21 < $C$24) * (C$11:C$21)))}`

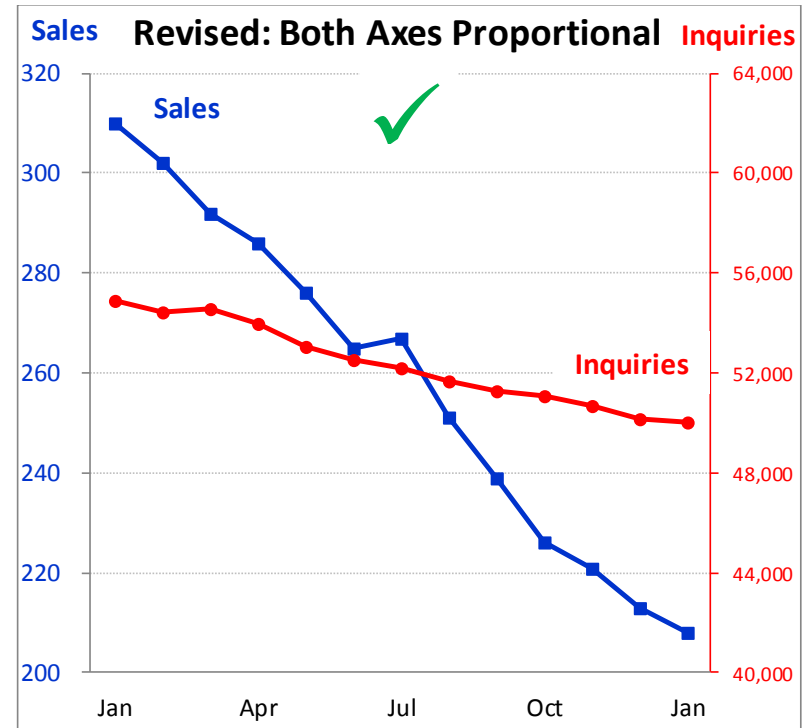
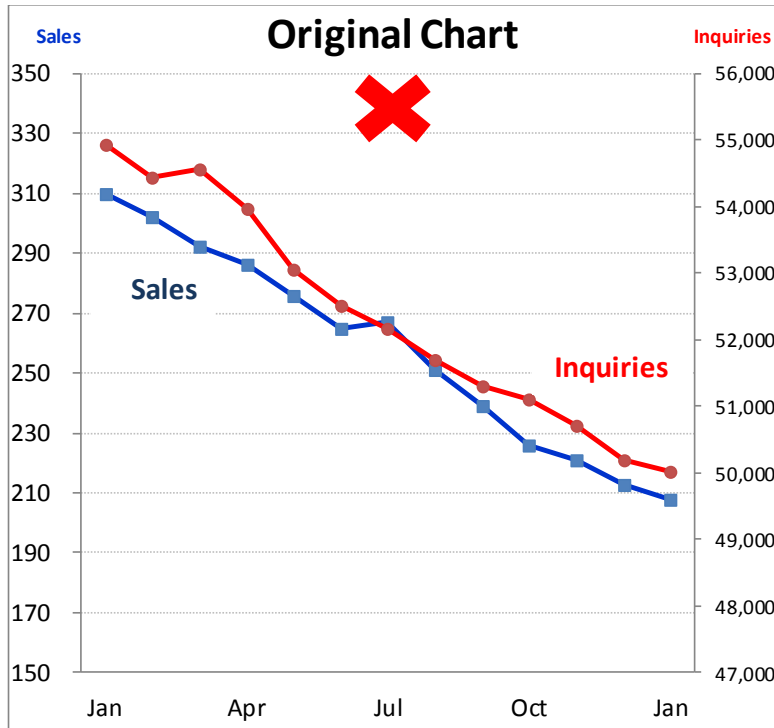
← Matches name  
 ← Sets 2<sup>nd</sup> criteria  
 ← Picks up Per. 1 values

Array formulas are another very powerful Excel feature. SUMIF & VLOOKUP formulas perform similar service, but they can't be as easily and universally employed as array formulas. I particularly like their flexibility – you can have one formula that sets one or more criteria, for different types of criteria, for an unsorted dataset, that's scattered across several worksheets. Once properly formulated, it can be easy to copy & edit, and becomes a very powerful tool. The simple example above has set two different criteria in the "Totals <100" section: that the name matches, & that it picks up Period values only when the "Total" (from Col. H) was under 100. Again, many criteria, of different types, can be applied in one formula.

Things to be aware of when using array formulas:

- The formula is invoked by holding down the Ctrl+Shift keys while hitting "Enter"; this gives it the funny brackets {}.
- When editing a formula, you'll always need to enter the change invoking "Ctrl+Shift+Enter".
- Array formulas can't be copied over themselves.

# 9. When inserting a 2<sup>nd</sup> Y-axis onto a chart, be sure it's proportionate with the primary axis.



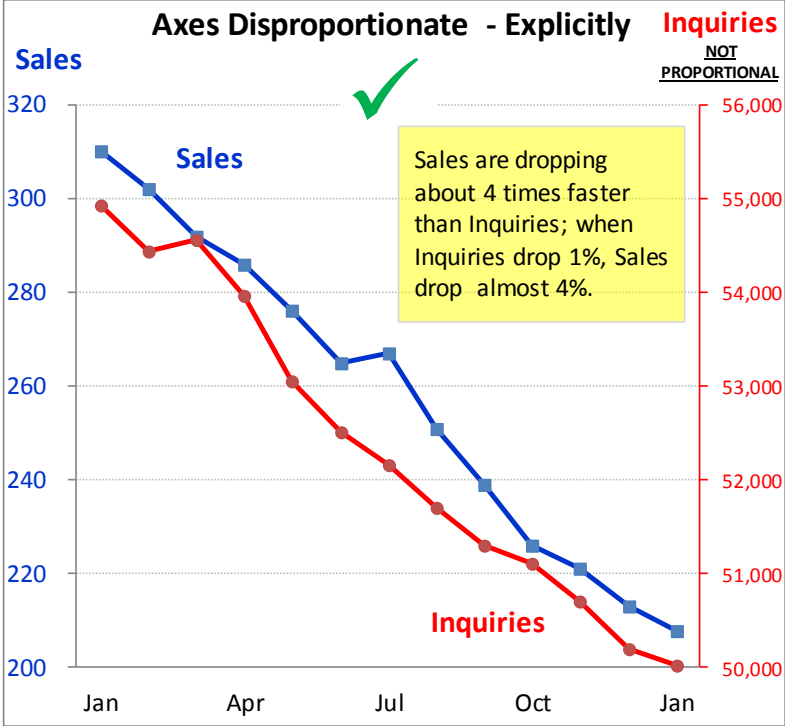
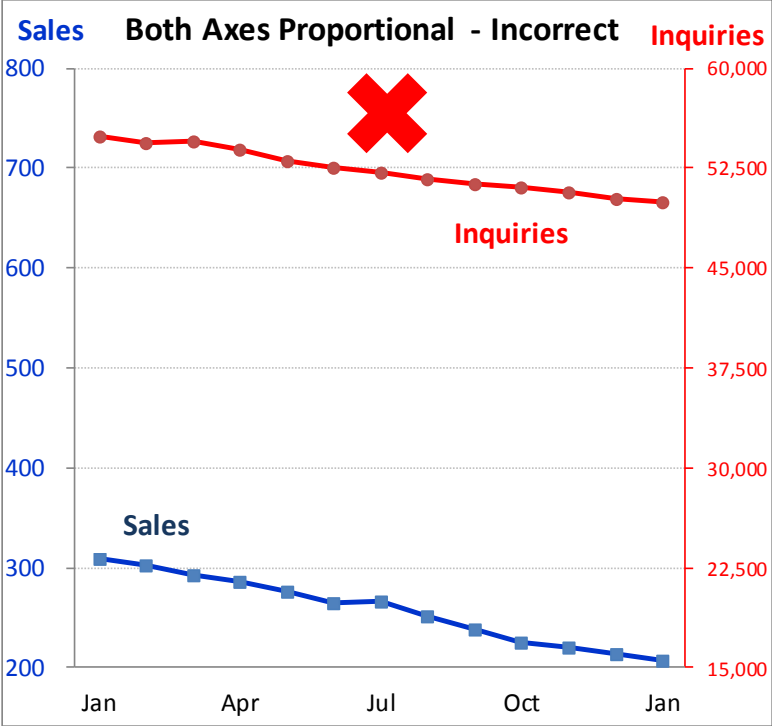
There are numerous occasions when charts are prepared that contain sets of data that are on very different scales. In the example at the left above, sales and inquiries are displayed. Because inquiries are so much greater in number, they are tracked on a 2<sup>nd</sup> y-axis, on the right side of the chart. The range displayed for the 2<sup>nd</sup> y-axis is

nicely set here such that the “Inquiries” line almost exactly parallels the “Sales” line. Such an appearance strongly implies that Sales & Inquiries are declining at the same pace. And that’s the problem.

The right-hand chart has been set such that the

2<sup>nd</sup> y-axis is exactly proportional with the 1st y-axis: by multiplying any number on the primary axis by 200, you get the figure shown on the 2<sup>nd</sup> y-axis. The revision clearly shows Sales dropping far more rapidly than Inquiries. This provides a much more accurate & valid comparison.

# Just forcing both axes to be proportional isn't enough. If necessary, be explicit that the two axes *aren't* proportional.

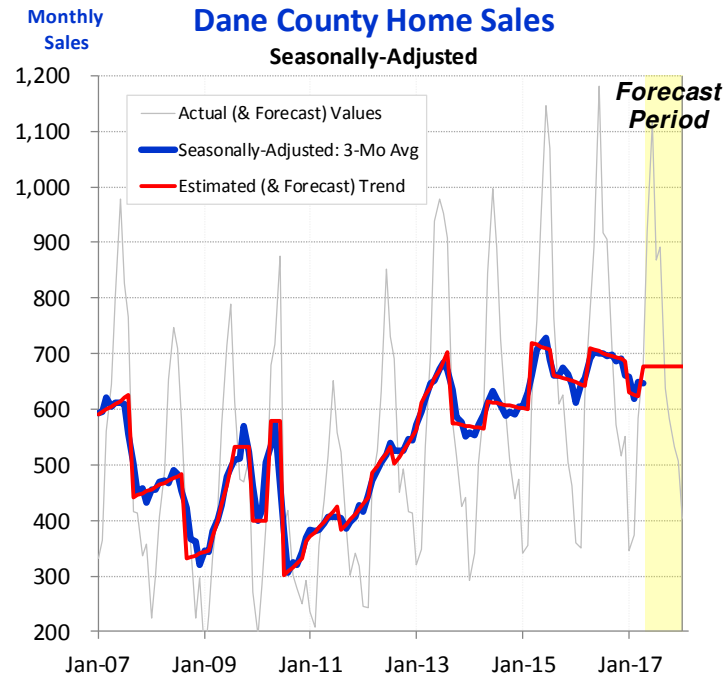


The above left-hand chart has both axes proportional: the right-hand axis is always 75x greater than the left-hand. Set up this way it again looks like Sales are declining at about the same pace as Inquiries; but they aren't. The deception is due to their relative placement against their axes – Sales are picking up at the bottom of their 200-800 range, while Inquiries are read at the top end of their 15,000-60,000 range. On a percentage basis, the Sales are dropping greater as they're at the bottom of their "range".

Another way to deal with this is to simply allow the two axes to NOT be proportional – you just have to spell it out, as is done here with the labeling at the top of the 2<sup>nd</sup> Y-axis, and with the inset Note describing how Sales are dropping about 4x faster than Inquiries.

# 10. It can be difficult to read a chart or slide if it has too much information.

After strong growth in 2011-13, sales have stabilized. We project modest growth, despite recent news reports.



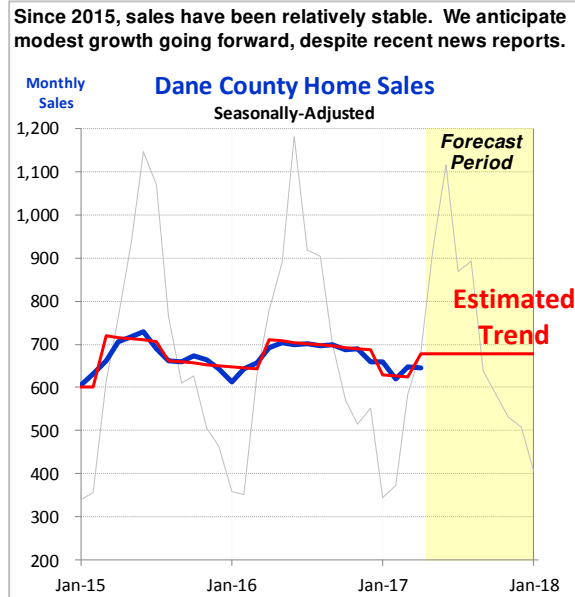
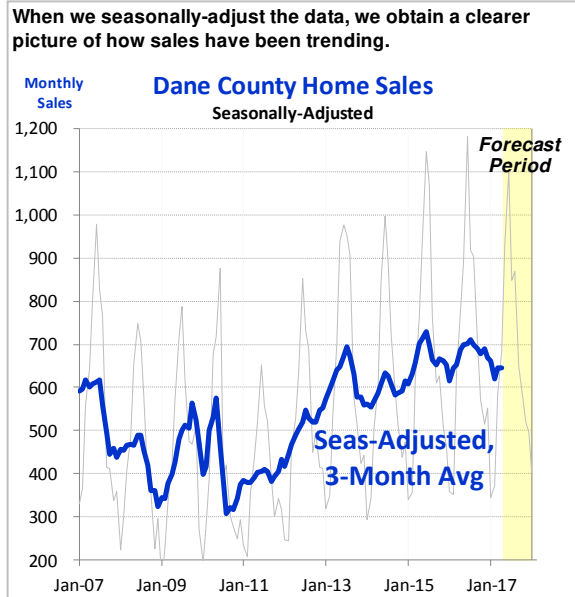
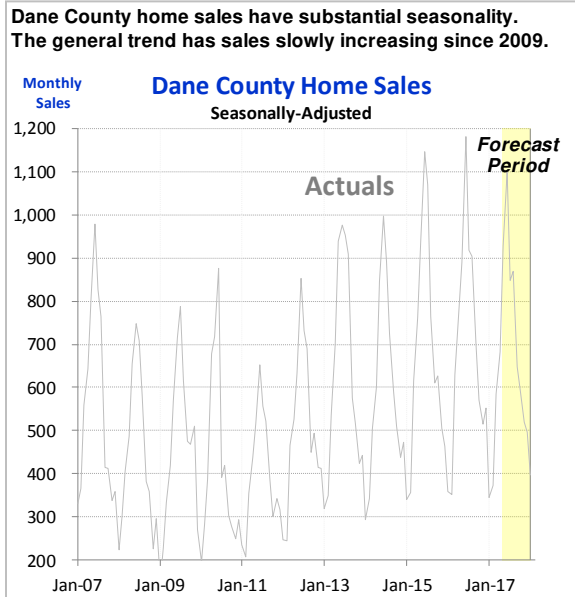
There is a natural tendency to squeeze lots of lines onto a single chart; and to present a full storyline on a single chart. But limiting the number of lines on a chart can be easier for a reader to follow; (again, less is often more). And splitting up a

storyline over a series of charts can be a more effective means of communicating.

This “slide” shown here uses a legend to identify the three lines. Without some explanation, it may

be difficult for the reader to follow what’s happening on this chart. Even if someone were describing it, they would have some difficulty pointing to the different lines to explain each one. And note this is a relatively “easy” example.

# Using a series of slides to tell a story can be much more effective.



The 3 “slides” shown here use the 3 pages to walk the reader through the 3 different chart lines. It makes it easier for the reader to follow along, and renders the final chart more meaningful.

For much excellent guidance on charts, I strongly recommend anything written by Stephen Few. He’s a master at focusing on showing only what is needed, and no more. For guidance on presenting a story, Cole Nussbaumer Knaflic’s

Storytelling with Data is superb – lots of good examples. I appreciated her emphasizing that the data work doesn’t end with the analysis, it’s the *communication* of the data that’s the true endpoint, and it takes time to do that well.