Using daily data in Tableau to calculate a seven-day moving average and moving range Daily Comp control chart pair, with control limits shown on each.

Written by Ken Black

This chart pair is typically called an x-bar and R chart. Descriptions of this chart an be found at: (http://en.wikipedia.org/wiki/Xbar_and_R_chart)

Step 1: Example data

Shown below is a sample of the data, with turnover being the total sales for a station per day. In Tableau, the turnover field is renamed Total Sales (€). This field is what is shown in the control charts. The database as 565 stations, about 500,000 rows of data from 2011-current, and is blended with another file to access geographical information for the stations.

	A	B	C	D	E	F	G	H	I
	Ship-to	party	Date	Site type	District	Sales Off.	Sales Grp	Pricing dt	turnover
2	XJ00004	MUENCHEN	01.01.2011	002		0003	006	1/1/2011	1,868.18
3	XJ00004	MUENCHEN	02.01.2011	002		0003	006	1/2/2011	2,797.67
4	XJ00004	MUENCHEN	03.01.2011	002		0003	006	1/3/2011	1,870.65
5	XJ00004	MUENCHEN	04.01.2011	002		0003	006	1/4/2011	1,697.22
6	XJ00004	MUENCHEN	05.01.2011	002		0003	006	1/5/2011	2,081.51
7	XJ00004	MUENCHEN	06.01.2011	002		0003	006	1/6/2011	2,397.09
8	XJ00004	MUENCHEN	07.01.2011	002		0003	006	1/7/2011	2,257.90
9	XJ00004	MUENCHEN	08.01.2011	002		0003	006	1/8/2011	2,112.48
10	XJ00004	MUENCHEN	09.01.2011	002		0003	006	1/9/2011	2,763.97
11	XJ00004	MUENCHEN	10.01.2011	002		0003	006	1/10/2011	2,440.65

Step 2: How to calculate a daily comparable on an annual basis

Let us use data from one store to demonstrate how to calculate the daily comparable on an annual basis. A daily comp is defined as:

(Total sales on a particular day in this year – Total sales on a particular day from last year)/ Total sales on a particular day from last year.

Example data:

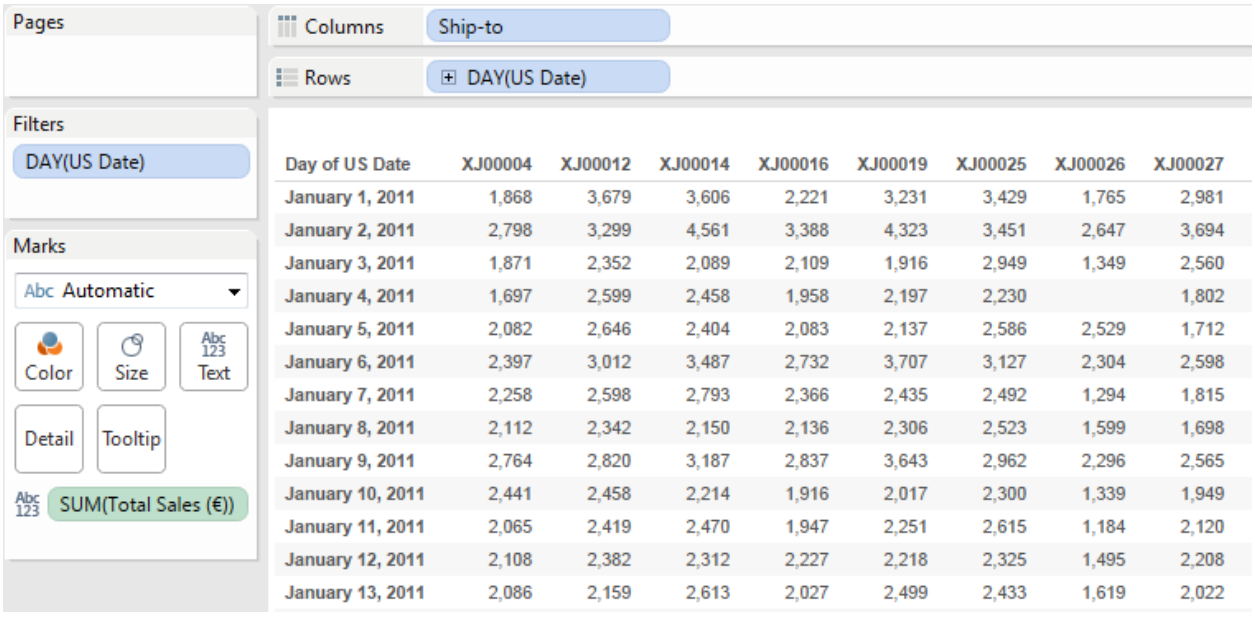
Total Sales at a Station for the same day in three consecutive years

Date	Total Sales	Annual Comparable Sales
Monday, January 03, 2011	1,870.7	
Monday, January 02, 2012	1,887.5	0.90%
Monday, December 31, 2012	2,243.2	18.85%

In this example, the day selected for this comparable was the first Monday in 2011, which happened to be on Jan 3, 2011. In the following year, the first Monday occurred on Jan 2, 2012. This one-day shift is normal except for leap years. Because 2012 was a leap year, there was a two-day shift in the third year data point which pushed the date back to December 31, 2012. All three of these days represent the same business day 1-year apart, which is why this comparable is called a daily comparable calculated on an annual basis. The 0.9% comparable for 1/2/2012 was calculated as (1887.5-1870.7)/1870.7. This means that sales in 2012 were about 1% better than the sales from the same day in 2011. One year after that, the sales improved nearly 19%. If the comparable is negative, it means that the sales in the previous year were better than the current year.

In this database, not every station has a daily reported sales total, so calculating a daily comp on an annual basis does present some challenges. Since each station has randomly missing data, if you are using Excel to hold your daily data, it is not possible to simply index the previous year daily comp sale by pointing to the row of data that is 365 (or 366 in the case of leap years) rows above the day in question. When you factor in the presence or absence of the Feb 29 leap day, it is tricky to make sure that the daily comp is being compared to the correct day from the previous year. The following step gives a simple solution to this problem rather than trying to write a custom computer program to align the data for this analysis.

Step 3: Pull the database into Tableau and write a properly structured Excel crosstab dataset of sales



Day of US Date	XJ00004	XJ00012	XJ00014	XJ00016	XJ00019	XJ00025	XJ00026	XJ00027
January 1, 2011	1,868	3,679	3,606	2,221	3,231	3,429	1,765	2,981
January 2, 2011	2,798	3,299	4,561	3,388	4,323	3,451	2,647	3,694
January 3, 2011	1,871	2,352	2,089	2,109	1,916	2,949	1,349	2,560
January 4, 2011	1,697	2,599	2,458	1,958	2,197	2,230		1,802
January 5, 2011	2,082	2,646	2,404	2,083	2,137	2,586	2,529	1,712
January 6, 2011	2,397	3,012	3,487	2,732	3,707	3,127	2,304	2,598
January 7, 2011	2,258	2,598	2,793	2,366	2,435	2,492	1,294	1,815
January 8, 2011	2,112	2,342	2,150	2,136	2,306	2,523	1,599	1,698
January 9, 2011	2,764	2,820	3,187	2,837	3,643	2,962	2,296	2,565
January 10, 2011	2,441	2,458	2,214	1,916	2,017	2,300	1,339	1,949
January 11, 2011	2,065	2,419	2,470	1,947	2,251	2,615	1,184	2,120
January 12, 2011	2,108	2,382	2,312	2,227	2,218	2,325	1,495	2,208
January 13, 2011	2,086	2,159	2,613	2,027	2,499	2,433	1,619	2,022

Note that the days are along the rows and the stations are along the columns. This table is then exported as a crosstab to Excel. The total size of this in excel is 852 rows by 584 columns as shown below (note that certain rows and columns are hidden in this view). The column B shown below is added to the table to count the days in each year to make sure that there is an entry for every day of the year since we are randomly missing data from each station. In this case, since there are over 580 stations in the database, all days of each year are accounted for.

	A	B	C	D	E	VK	VL
1	Date	Day of Year	XJ00004	XJ00012	XJ00014	XJ05843	XJ05856
2	Saturday, January 01, 2011	1	1868.18	3678.96	3606.01		
3	Sunday, January 02, 2011	2	2797.67	3298.66	4561.41		
4	Monday, January 03, 2011	3	1870.65	2351.69	2089.05		
364	Thursday, December 29, 2011	363	1776.22	2369.88	2588.59		
365	Friday, December 30, 2011	364	2394.23	2641.56	2669.98		
366	Saturday, December 31, 2011	365	2638.17	2975.76	4222.58		
367	Sunday, January 01, 2012	1	2181.00	3046.85	3025.40		
368	Monday, January 02, 2012	2	1887.52	2412.67	2135.21		
369	Tuesday, January 03, 2012	3	1801.39	2250.33	1951.38		
729	Friday, December 28, 2012	363	2117.93	2449.14	2608.48		745.18
730	Saturday, December 29, 2012	364	2960.12	2249.24	2304.03		1027.12
731	Sunday, December 30, 2012	365	2598.10	2445.45	3500.36		1159.75
732	Monday, December 31, 2012	366	2243.23	1918.97	2782.29		1110.60
733	Tuesday, January 01, 2013	1	1824.06	2436.61	2810.23		971.38
734	Wednesday, January 02, 2013	2	1765.05	2495.96	2420.79		1009.33
735	Thursday, January 03, 2013	3	1785.15	2028.76	2562.45		695.84
850	Sunday, April 28, 2013	118	3016.85	2115.61	3247.29	1675.22	2093.07
851	Monday, April 29, 2013	119	1700.05	1599.72	2579.19	1653.49	1433.29
852	Tuesday, April 30, 2013	120	2329.79	2081.34	3193.36	2081.98	1767.59
853	1	2	3	4	5	583	584

Step 4: Calculate the daily comps for each store in the database

Using the sales information stored in the the table from Step 3 above, the daily comparables can be calculated using the formulas shown below. Note that the first comparable formula shown in the table below in Cell C2 is for Sunday, Jan 1, 2012, which uses the values in rows 3 (Sun Jan 2, 2011) and 367 (Sun Jan 1, 2012) of the table above in Step 3. It is very important to get the alignment correct so that each daily comparable is being compared to the same day the year before. Because the master table from step 3 has every day of the year in it, the formula only has to be written once in Cell C2 and then can be copied across the entire table. In the second table below, you can see the calculated daily comparables for each station for each day. Although this crosstab view could be reimported to Tableau to create the control charts, this data structure is inefficient and would create over 580 individual measures, which is not what we need to complete our control charts. One additional step is required to make our analysis much easier.

	A	B	C	D
1	Date	Day of Year	XJ00004	XJ0001
2	40909	1	=+(Total_Sales!C367-Total_Sales!C3) / Total_Sales!C3	=+(Total_Sales!D367-Total_Sales!D3)/Total_Sales!D3
3	40910	=+B2+1	=+(Total_Sales!C368-Total_Sales!C4)/Total_Sales!C4	=+(Total_Sales!D368-Total_Sales!D4)/Total_Sales!D4
4	40911	=+B3+1	=+(Total_Sales!C369-Total_Sales!C5)/Total_Sales!C5	=+(Total_Sales!D369-Total_Sales!D5)/Total_Sales!D5
5	40912	=+B4+1	=+(Total_Sales!C370-Total_Sales!C6)/Total_Sales!C6	=+(Total_Sales!D370-Total_Sales!D6)/Total_Sales!D6
6	40913	=+B5+1	=+(Total_Sales!C371-Total_Sales!C7)/Total_Sales!C7	=+(Total_Sales!D371-Total_Sales!D7)/Total_Sales!D7
220	41127	=+B219+1	=+(Total_Sales!C585-Total_Sales!C221)/Total_Sales!C221	=+(Total_Sales!D585-Total_Sales!D221)/Total_Sales!D221
221	41128	=+B220+1	=+(Total_Sales!C586-Total_Sales!C222)/Total_Sales!C222	=+(Total_Sales!D586-Total_Sales!D222)/Total_Sales!D222
241	41148	=+B240+1	=+(Total_Sales!C606-Total_Sales!C242)/Total_Sales!C242	=+(Total_Sales!D606-Total_Sales!D242)/Total_Sales!D242
242	41149	=+B241+1	=+(Total_Sales!C607-Total_Sales!C243)/Total_Sales!C243	=+(Total_Sales!D607-Total_Sales!D243)/Total_Sales!D243

	A	B	C	D	E	F	G	H	I
1	Date	Day of Year	XJ00004	XJ00012	XJ00014	XJ00016	XJ00019	XJ00025	XJ00026
2	Sunday, January 01, 2012	1	-22.04%	-7.63%	-33.67%	-33.11%	-25.67%	-0.76%	-25.22%
3	Monday, January 02, 2012	2	0.90%	2.59%	2.21%	2.73%	35.70%	-14.88%	-0.10%
4	Tuesday, January 03, 2012	3	6.14%	-13.42%	-20.61%	3.34%	18.28%	17.33%	#DIV/0!
5	Wednesday, January 04, 2012	4	-15.15%	-11.09%	0.36%	-4.86%	1.62%	-4.51%	-44.13%
6	Thursday, January 05, 2012	5	-2.28%	-13.65%	-28.22%	0.29%	-29.82%	-9.79%	-43.22%
220	Monday, August 06, 2012	219	-6.37%	22.24%	-10.55%	17.05%	#DIV/0!	-6.70%	-14.98%
221	Tuesday, August 07, 2012	220	-20.69%	9.40%	2.24%	26.50%	-45.12%	-7.62%	-15.02%
241	Monday, August 27, 2012	240	-24.12%	13.13%	-15.58%	20.05%	-85.57%	20.85%	-17.82%
242	Tuesday, August 28, 2012	241	11.84%	-14.16%	-1.78%	-8.89%	-100.00%	-2.38%	2.23%
243	Wednesday, August 29, 2012	242	3.85%	-4.04%	-11.61%	-5.89%	-100.00%	1.74%	-2.24%
244	Thursday, August 30, 2012	243	3.20%	-8.89%	-4.38%	8.79%	-100.00%	5.99%	-9.69%

Note that there are two types of data problems in the daily comp table shown above. The first issue is #Div/0!. This occurs when there is missing daily data from the previous year (see cells I4 and G220). The second issue is -100.00%, which occurs when there is missing daily data from the current year (see cells G242 to G244). These values can simply be removed in Tableau via filtering in the next step.

Step 5: Restructure the comp data so it can be imported to Tableau to generate the control charts

Using the table of daily comps created in Step 4, you can reshape the data into four columns so that it can be imported to Tableau. This process is documented in a video: Click here to watch the reshaping data video.

Step 6: Create the control charts

With the data from step 5 imported to Tableau, the control charts can be created. This process is documented in a video: Click here to watch the creating control charts video.

A lot of the information contained in the creating control charts video is documented in a web blog that can be accessed by clicking here.