Building dynamic process control charts with SAS ® Visual Analytics in Pharma and Healthcare

Massimo Fabris, SAS Italy

ABSTRACT

Shewhart control chart are used as a monitoring tool in pharmaceutical manufacturing process to reduce the reject ratio and promote the continuous improvement of products. As an example, control charts are used in Pharma for Product Quality Review (PQR) which is a clear requirement of all cGMP guidelines and is a powerful tool for measuring and optimizing not only product quality but also process efficiency. In Healthcare, the purpose of statistical process control (SPC) is often to quantify improvements and identify unintended consequences resulting from an intentional change in an environment, policy, treatment protocol.

In this paper we demonstrate how SAS® Visual Analytics can be used to build dynamic and interactive XS control chart. The analyst can perform actions such as the following: defining a baseline period and extending its center line, UCL and LCL prospectively; removing any subgroup from baseline calculations; selecting control chart which violates limits and western electric rules; adjusting the sigma multiplier. The same methodology can be applied to build other Shewhart control charts.

This approach enhances and complements a traditional batch approach based on SAS/QC® by providing a high degree of interactivity and flexibility on the data visualization phase.

INTRODUCTION

This article illustrates how to use with SAS Visual Analytics to create one of the most popular control charts in Statistical Process Control (SPC). Before going on with the illustration of the control chart, some fundamental concepts are briefly introduced.

Statistical Process Control is a method of quality control which employs statistical methods to measure, monitor, and control a process.

Control charts are ones of the primary techniques of statistical process control. A control chart is a graphical display of quality characteristics that have been measured versus the sample number or production time. The control chart was invented by Walter Shewhart at Bell labs in 1920.

The control chart is hence a graph used to study how process changes over time. A control chart usually has a central line for the (estimated) average of the process measure, an upper line for upper control limit, and lower line for the lower control limit. The control limits are usually ±3 (estimated) standard deviation of the measure from the centerline.

We can say a process is out-of-control if any point in the control chart is out of control limits or having abnormal patterns of variability. Teams must identify special causes and try to eliminate to achieve a stable process (Hessing, Ted. 2022).

The aim of this paper is describing a real-world use case requirement for a flexible and interactive control chart which arise in the Product Quality Review of a Pharma production process and how this requirement can be accomplished with SAS Visual Analytics.

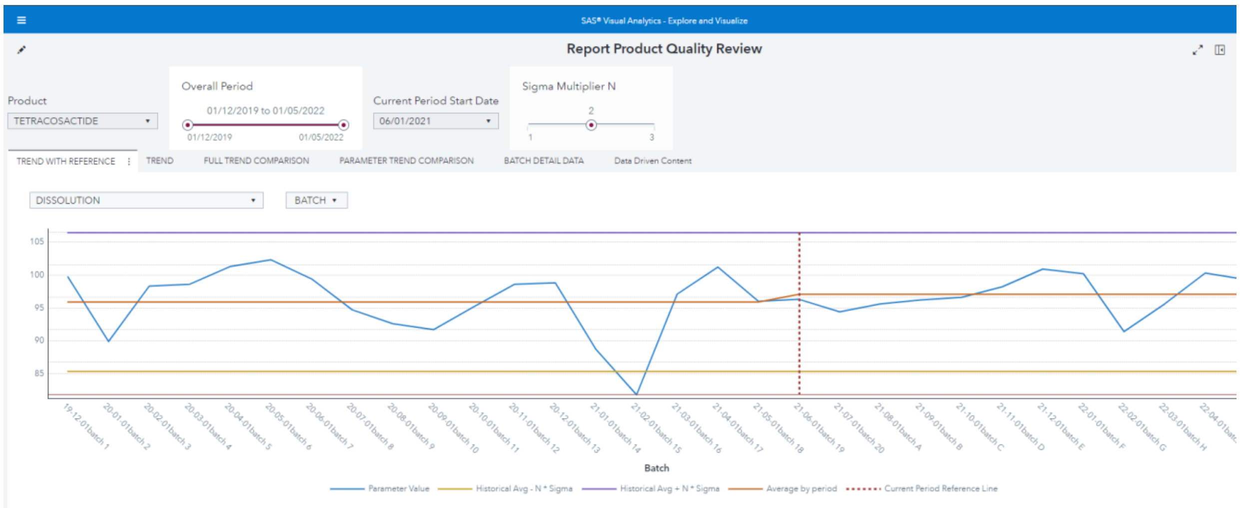
The paper is organized as follows: the first section illustrates the requirements from the real-world use case and it shows how such dynamic chart can be realized with SAS Visual Analytics. The second section introduces motivations and benefits for this flexibility requirements by providing literature references. The third section shows how the use case can be generalized to produce a chart which respects the specification of the XS chart of SAS QC proc SHEWHART. The fourth section discusses how much this approach can be extended to the complete taxonomy of control charts. The fifth section advocates the benefits of implementing control charts with SAS Visual Analytics by itself or by enhancing and complementing the traditional approach based on proc Shewhart in SAS QC or proc SPC in SAS Visual Statistics.

REAL WORLD USE CASE REQUIREMENT

Recently, two different Pharma companies have expressed the need to provide their analysts with a highly dynamic control chart. The requirement is to provide the user with high flexibility in selecting batches and time periods. The user should hence be allowed to define the date splitting the time period in two parts: one historical and one current. The report should hence calculate central line and control limits on the historical period values and compare them with current period values and current central line.

Before illustrating in the next section the motivations and benefits for this flexibility requirements, we illustrate how all these requirements were accomplished and how the user can interact with the analysis.

Figure 1 Product Quality Review control chartshows the SAS Visual Analytics report which fulfills all requirements mentioned above:



Product	Param Name	Batch Date	Batch N.	Parameter Value	Average by period	Historical Avg - N * Sigma	Historical Avg + N * Sigma	Min	Max	Batch
TETRACOSACTIDE	DISSOLUTION	01/10/2020	11	95.20	95.88	85.34	106.42	81.80	102.30	20-10-01batch 11
TETRACOSACTIDE	DISSOLUTION	01/11/2020	12	98.60	95.88	85.34	106.42	81.80	102.30	20-11-01batch 12
TETRACOSACTIDE	DISSOLUTION	01/12/2020	13	98.80	95.88	85.34	106.42	81.80	102.30	20-12-01batch 13
TETRACOSACTIDE	DISSOLUTION	01/01/2021	14	88.70	95.88	85.34	106.42	81.80	102.30	21-01-01batch 14
TETRACOSACTIDE	DISSOLUTION	01/02/2021	15	81.80	95.88	85.34	106.42	81.80	102.30	21-02-01batch 15
TETRACOSACTIDE	DISSOLUTION	01/03/2021	16	97.10	95.88	85.34	106.42	81.80	102.30	21-03-01batch 16

Figure 1 Product Quality Review control chart

The user analyst can select from left to right and from top to bottom:

1. The product.
2. The overall period.
3. How to split the overall time period in two part, historical and current, by selecting the "Current Period Start Date".
4. The number of sigma multiplier N to use for the control limits definition.
5. The measure (parameter) to plot.
6. Which batches to include/exclude (also by selecting batch properties).

Note that:

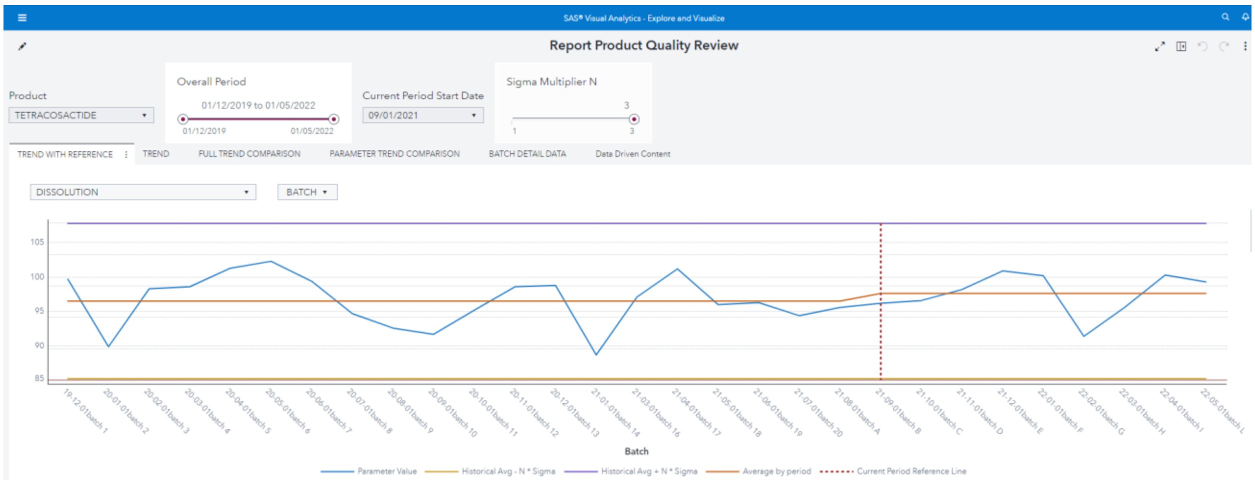
- The "current period" red dotted reference line is dynamic and added with a Custom Graph by using a Line chart and a Needle chart.
- The "Average by period" line calculates the average in two different periods so it is not a straight line.

- The batch n. 15 violates the Lower Control Limit.

As an example of the required flexibility, the user can:

- Remove the batch n. 15 because s/he knows there was an exceptional issue in production.
- Change The “current start date” to 1 Sep 2021.
- Augment the sigma multiplier N to 3.

The chart immediately reflects user inputs:



In the next paragraphs, we are going to show how the above control chart can be implemented.

BASIC INPUT TABLE DATA STRUCTURE

Table 1 depicts the required basic data structure.

FIELD	DESCRIPTION
Batch N	Batch identifier
Batch Date	Usually batch production date
Param Name	Measure Name
Param Value	Measure Value

Table 1 - Basic data structure

Figure 2 shows an example of the basic input table. The table can have any number of additional classifiers which may be used to filter batches, periods and parameters.

Batch N.	Batch Date ▲	Param Name ▲	Parameter Value
1	01/12/2019	CONTENUTO_IN_ACQUA	3.63
1	01/12/2019	DISSOLUTION	99.70
1	01/12/2019	DUREZZA	7.20
1	01/12/2019	PESO_MEDIO	622.20
1	01/12/2019	TITOLO_ENZIMATICO_POST_PACK	316.70
1	01/12/2019	TITOLO_ENZIMATICO_PRE_PACK	326.00
1	01/12/2019	UNIFOR_DOSE_UNITARIA_EP	2.20
2	01/01/2020	CONTENUTO_IN_ACQUA	4.49

Figure 2 – Basic Input Table example

CALCULATED ITEMS IN SAS VISUAL ANALYTICS

Whatever parameter the user selects, the first step is distinguishing between Historical Parameter values and Current Parameter values.

The Historical Parameter value considers only parameter values of batches produced before the reference date of the Current Period Start Date:

```
IF ( Batch Production Date < Current Period Start Date )
RETURN Parameter Value
ELSE Missing
```

Figure 3 Historical Parameter definition

Obviously, the Current Parameter is the above formula with “Batch Production Date” greater or equal than “Current Period Start Date”

The Historical Avg is calculated as the average of Historical Parameter:

```
Avg _ForAll_ ( Historical Paramameter )
```

Figure 4 - Historical Avg definition

Note that:

- The function Avg considers only not missing data and hence only historical data.

- the _ForAll_ operator computes the average of all historical values currently selected in the overall period as required. If you use the default _ByGroup_ operator, the average is calculated for each day and this does not conform to requirements.

The last step is to compute Lower Control Limit (LCL) and Upper Control Limit (UCL). The LCL definition requested by the customer is:

Figure 5 - Lower Control Limit definition

Note, once again, that by using the _ForAll_ operator, the result is calculated aggregating all dates included in the overall period, hence the UCL and LCL produce a straigh line because their value is the same for all days included in the overall period.

The last requirement is to plot the union line of the historical and the current values. This is accomplished as below:

Figure 6 - Average by Period definition

MOTIVATIONS FOR THE ANALYSIS AND REFERENCE LITERATURE

Motivations of the aforementioned flexibility are explained by R. Jones and L. Liu in "Leveraging SHEWHART Procedure Options to Monitor and Evaluate Improvements in Healthcare" (2018). Authors state:

> In healthcare, the purpose of statistical process control (SPC) is often to quantify improvements and identify unintended consequences resulting from an intentional change in an environment, policy, treatment protocol, or decision-support tool […]
> Unlike in manufacturing, process change - rather than stability - is commonly sought, and interventions might be frequent and staggered over time.
> (Jones and Liu 2018, 1).

Hence not only process stability but also effects in intentional changes can be monitored with control charts. In the same article we read:

> *The Health Care Data Guide: Learning from Data for Improvement* (Provost and Murray 2011) provides guidance to healthcare institutions to develop and operationalize the knowledge and tools needed for "Learning from Variation in Data" through the use of SPC. The section titled "Establishing and Revising Limits for Shewhart Charts" highlights several nuances operational leaders and their statisticians should understand and consider in analyzing the impact of intervention. These include defining a baseline period; "ghosting" data points in the baseline period that are considered special cause variation; and detecting improvements after freezing and extending an initial mean, or center line (Jones and Liu 2018, 1).

Control charts are not used only as a passive tool to monitor stability, but also as an active and dynamic tool:

> The main purpose for defining the center line and control limits for the baseline period is the generation of a comparison period against which improvements or unintended consequences of the intervention can be detected and statistically quantified (Jones and Liu 2018, 1).

A MORE GENERIC APPROACH

The question then arises: how far can we go in replicating with SAS Visual Analytics the solid and exhaustive procedures like proc Shewhart in SAS QC and proc SPC in SAS Visual Statistics?

The goal is to replicate in SAS Visual Analytics as much as possible of the control chart taxonomy:

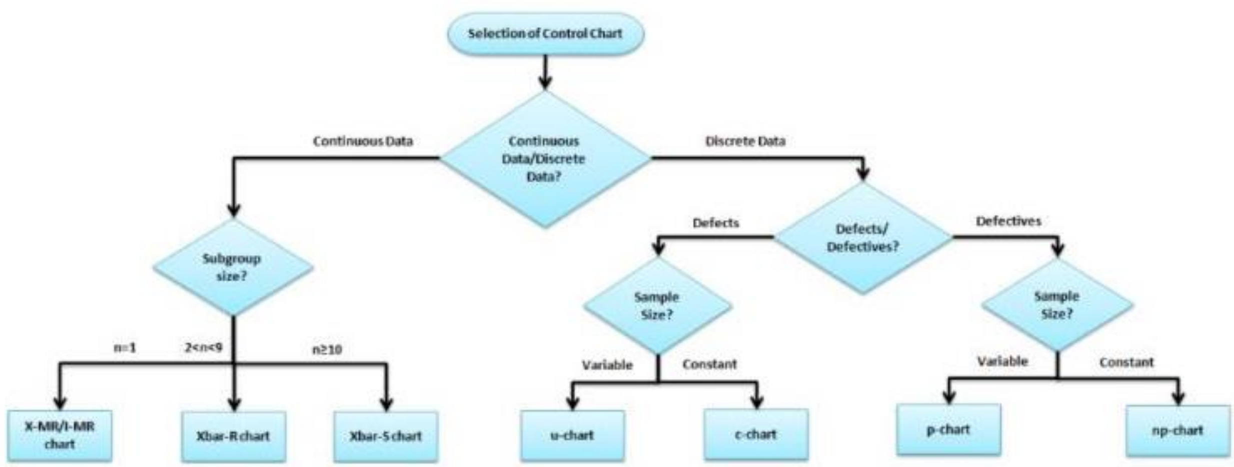


Figure 7 - Process Control Chart Taxonomy

We start with XS chart (Xbar-S chart), given its popularity among the control charts.

IMPLEMENTING PROC SHEWHART XS CHART IN SAS VISUAL ANALYTICS

We now show how to replicate the XS Chart of proc Shewhart which corresponds to the Xbar-S chart (Hessing, Ted 2022) in the taxonomy above. The general case of N samples (subgroups) with different sizes is considered and the proc Shewhart diagram is replicated with SAS Visual Analytics.

In the following Image, we have the XS chart built with SAS/QC proc Shewhart. The X chart is at the top and plots the average value of each subgroup. The S chart is at the bottom and plots the standard deviation of each subgroup.

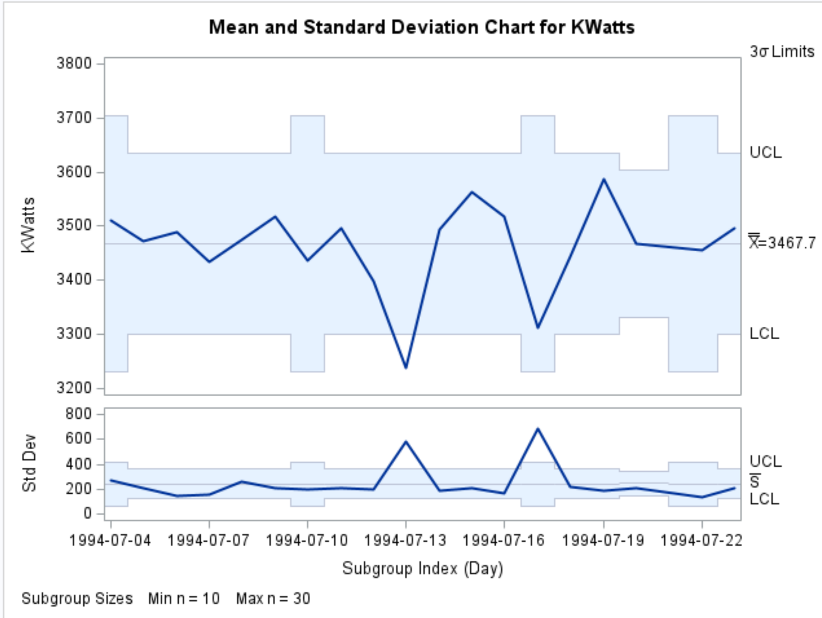


Figure 8 - XS chart with SAS QC proc Shewhart

Below is the corresponding graph on SAS Visual Analytic. Note that only the X chart is represented (but the S chart may be done similarly):

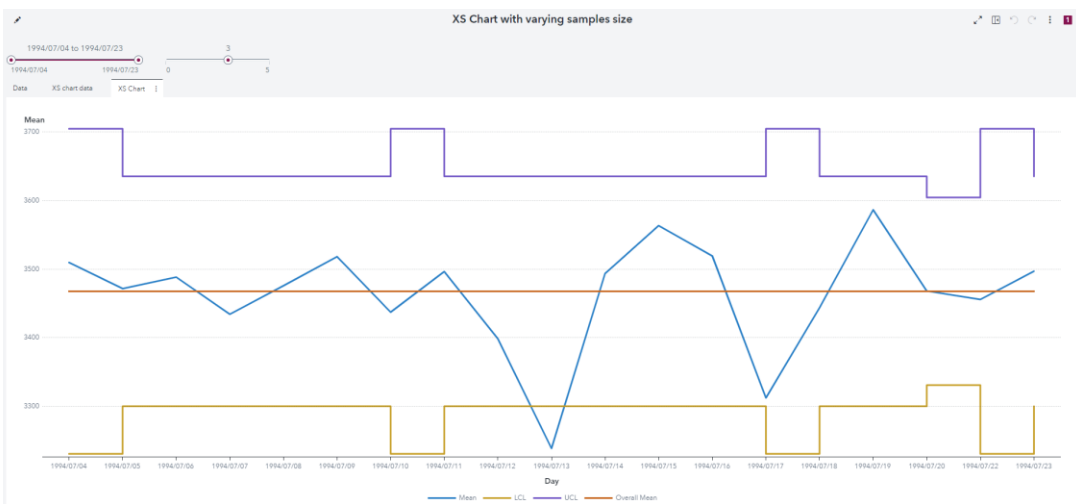


Figure 9 - X Chart with SAS Visual Analytics

The graph above is a SAS Visual Analytics Custom Graph combining together a Line chart with a Step Plot chart. This is an option to consider if in need of plotting a control chart with different subgroup size and you do not want to use the standard line graphs object.

In the next paragraphs we are going to describe how to produce this graph.

INPUT DATA

The input data set is made by 19 days of measurements of the amount of power (in kilowatts) used to heat the water to the desired temperature for a production process. Each day has a different number n(i) of measurements.

The SAS System

Obs	Day	KWatts
1	1994-07-04	3196
2	1994-07-04	3507
3	1994-07-04	4050
4	1994-07-04	3215
5	1994-07-04	3583

Figure 10 - Input data table

Each day represents a different subgroup in the XS chart in Figure 8. KWatts is the so called "process variable" the XS chart is meant to monitor.

XS CHART BACKGROUND DETAILS

We refer to "*SAS/QC® 15.2 User's Guide*" issued by SAS Institute in 2020 (SAS Institute Inc. 2020, 1990) for the notations and formulas:

$\bar{X}_i$ Mean of measurements in ith subgroup

s_i Standard deviation of the measurements $x_{i1}, \ldots, x_{in_i}$ in the ith subgroup

$$s_i = \sqrt{((x_{i1} - \bar{X}_i)^2 + \cdots + (x_{in_i} - \bar{X}_i)^2)/(n_i - 1)}$$

n_i Sample size of ith subgroup

N Number of subgroups

$\bar{\bar{X}}$ Weighted average of subgroup means

Figure 11 - Notations and formulas

The two key calculations on a XS chart are the estimation of the Mean and the Standard Deviation of the (process) measures.

The Mean estimate is used as central line in the X chart and is calculated as a *weighted average of the subgroups means* (SAS Institute Inc. 2020, 1990).

Central Lines

On an $\bar{X}$ chart, by default, the central line indicates an estimate of μ, which is computed as

$$\hat{\mu} = \bar{\bar{X}} = \frac{n_1\bar{X}_1 + \cdots + n_N\bar{X}_N}{n_1 + \cdots + n_N}$$

If you specify a known value (μ_0) for μ, the central line indicates the value of μ_0.

There are several methods to estimate the Standard Deviation. The default method used by proc Shewhart is to compute the average of the estimated standard deviations of each subgroup. The estimation of the standard deviation on each subgroup is calculated as the standard deviation of the values in the (i) subgroup divided by the factor C4 which depends on the size of the subgroup n(i). This factor can be calculated using the C4(X) function (SAS Institute Inc. 2020, 2272).

A more complete formalization of the Standard Deviation estimate is the following (SAS Institute Inc. 2020, 1497):

Default Method Based on Subgroup Standard Deviations

If you do not specify the RANGES option, the default estimate for σ is

$$\hat{\sigma} = \frac{s_1/c_4(n_1) + \cdots + s_N/c_4(n_N)}{N}$$

where N is the number of subgroups for which $n_i \geq 2$, s_i is the sample standard deviation of the ith subgroup

$$s_i = \sqrt{\frac{1}{n_i - 1}\sum_{j=1}^{n_i}(x_{ij} - \bar{X}_i)^2}$$

and

$$c_4(n_i) = \frac{\Gamma(n_i/2)\sqrt{2/(n_i - 1)}}{\Gamma((n_i - 1)/2)}$$

Here $\Gamma(\cdot)$ denotes the gamma function, and $\bar{X}_i$ denotes the ith subgroup mean. A subgroup standard deviation s_i is included in the calculation only if $n_i \geq 2$. If the observations are normally distributed, the expected value of s_i is $c_4(n_i)\sigma$. Thus, $\hat{\sigma}$ is the unweighted average of N unbiased estimates of σ. This method is described in the American Society for Testing and Materials (1976).

Figure 12 - Standard Deviation Estimate

XS CHART CALCULATED IN SAS VISUAL ANALYTICS

Now that all formulas are defined, we can start calculating: the first step is data pre-processing in SAS Base to calculate all the basic components. We calculate N (sample size), Mean, Std Dev for each subgroup (each day) using SAS Base proc means.

```
proc means data=TurbineDataWithVaryingSubgroupSize nway missing;
class Day;
var kwatts;
output out=StatByDay(drop=_type_ _freq_) n=n mean=mean std=std ;
run;
```

Then we associate to each group the corresponding C4 value which can be calculated or imported from a C4 table listing (you do not need to compute C4 factors if the subgroup size is predictable and limited – the are many sources of such table listing).

This is the final table to load on SAS Visual Analytics:

Obs	Day	n	mean	std	C4
1	1994-07-04	10	3509.60	275.663	0.97266
2	1994-07-05	20	3471.65	210.427	0.98693
3	1994-07-06	20	3488.30	147.025	0.98693
4	1994-07-07	20	3434.20	157.637	0.98693
5	1994-07-08	20	3475.80	258.949	0.98693

Figure 13 - Table ready for SAS Visual Analytics

Please note that:

- This aggregation step does not affect the correctness of subsequent calculations of the estimated mean, standard deviation, LCL and UCL in SAS Visual Analytics whatever is the set of subgroups the user selects in the report.
- The above data structure can be generalized to contain any number of batch production data with any desired batch classifier (product, plant, machine, faulty batch,…).

At this point we let SAS Visual Analytics do at run time the remaining calculations which depend on the user's filters:

(Sum _ForAll_ ((Mean * N)) / Sum _ForAll_ (N))

Figure 14 - Weighted average of subgroup means

Avg _ForAll_ ((StdDev / C4))

Figure 15 - Estimation of the Standard Deviation

Sum _ByGroup_ (N)

Figure 16 - Sample size

Now we can finally compute the Upper Control Limits (UCL) and Lower Control Limits (LCL) as they are defined in the user's guide mentioned above (SAS Institute Inc. 2020, 1991).

Control Limits

$\bar{X}$ Chart

$$\text{LCL} = \text{lower limit} = \overline{\overline{X}} - k\hat{\sigma}/\sqrt{n_i}$$

$$\text{UCL} = \text{upper limit} = \overline{\overline{X}} + k\hat{\sigma}/\sqrt{n_i}$$

Figure 17 - Control Limits definition

Note that:

- Since subgroups have different cardinality, the n(i) values are different. Hence the UCL and LCL are different for each subgroup and hence the UCL and LCL lines are not straight lines
- In the special case in which all subgroups have the same cardinality n, UCL and LCL are all equals for all subgroups and hence the UCL and LCL lines are straight lines.

This is the corresponding calculation in SAS Visual Analytics for UCL:

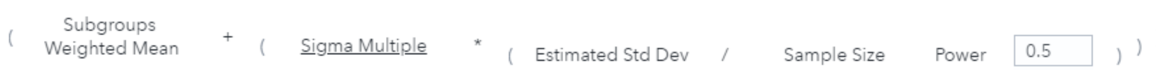
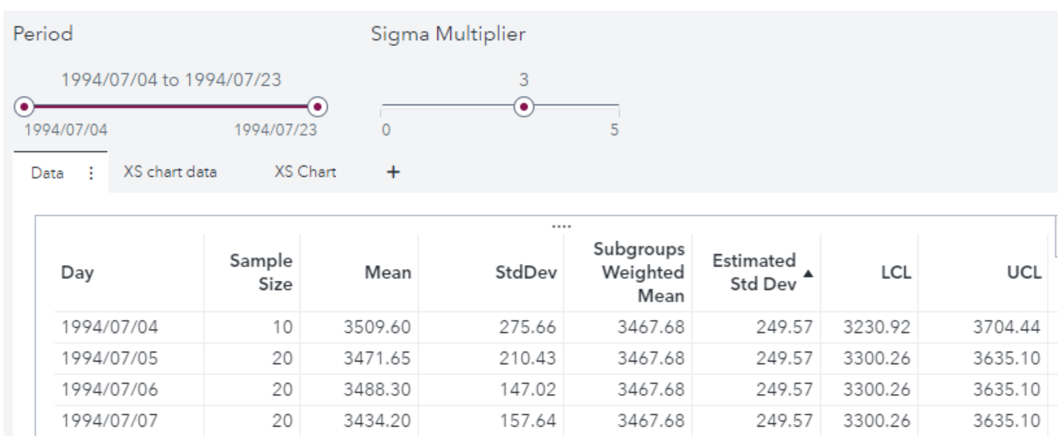


Figure 18 - UCL definition in SAS VA

Note that *k* corresponds to the user parameter "Sigma Multiple".

Below is a listing of the obtained data:



Day	Sample Size	Mean	StdDev	Subgroups Weighted Mean	Estimated Std Dev	LCL	UCL
1994/07/04	10	3509.60	275.66	3467.68	249.57	3230.92	3704.44
1994/07/05	20	3471.65	210.43	3467.68	249.57	3300.26	3635.10
1994/07/06	20	3488.30	147.02	3467.68	249.57	3300.26	3635.10
1994/07/07	20	3434.20	157.64	3467.68	249.57	3300.26	3635.10

The values are identical to the results obtained with proc Shewhart:

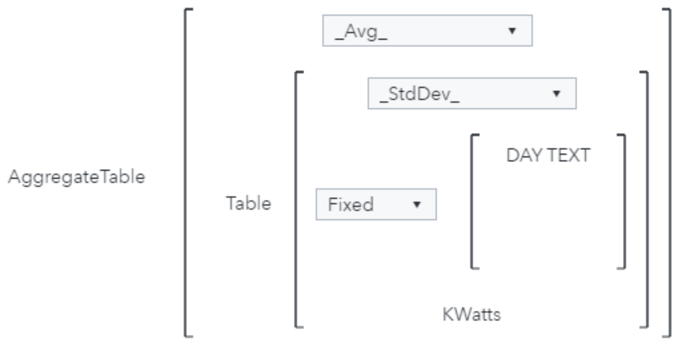
The SAS System

Obs	Day	Subgroup Sample Size	Subgroup Mean	Subgroup Standard Deviation	Process Mean	Process Standard Deviation	Lower Control Limit for Mean	Upper Control Limit for Mean
1	1994-07-04	10	3509.60	275.663	3467.68	249.568	3230.92	3704.44
2	1994-07-05	20	3471.65	210.427	3467.68	249.568	3300.26	3635.10
3	1994-07-06	20	3488.30	147.025	3467.68	249.568	3300.26	3635.10
4	1994-07-07	20	3434.20	157.637	3467.68	249.568	3300.26	3635.10

The appendix includes a program to generate proc Shewhart chart and compare it with SAS Base calculations.

SPECIAL CASE

It is worth mentioning that if all subgroups have the same cardinality, it is possible to avoid the data preprocessing proc means step to obtain the input table listed in Figure 13. This can be achieved by using the AggregateTable operator to compute the average of the standard deviations:



However this approach is limited to this special case and all other formulas are more complex. There are no evident advantages in working with detail data.

CONTROL CHAR TAXONOMY COVERAGE

This is the list of charts included in the taxonomy of Figure 7 and available in proc Shewhart. They can all be implemented with SAS Visual Analytics.

Proc Shewhart chart type	**Replicable in SAS Visual Anaytics**	**Notes**
IR chart	Yes*	With limitations in filtering
M chart	Yes	Using median, estimated std, and control chart constants
MR chart	Yes	Using median, ranges, estimated std, and control chart constants
XR chart	Yes	Using mean, ranges, estimated std, and control chart constants
XS chart	Yes	As shown in the paper
U chart	Yes	
C chart	Yes	
P chart	Yes	
NP chart	Yes	

COMPLEMENTING AND ENHANCING BATCH PROCESSING

The approach described in this paper can be implemented with SAS Visual Analytics alone, but it can benefit, complement and enhance the results obtained with proc Shewhart and proc SPC.

Both proc Shewhart and proc SPC can:

- Export the full detail of the results.
- Export summary tables with statistics on limits violation.
- Test for presence of special patters as "western electric rules".

As an example of these functionalities, let us consider the default output from PROC SPC is an exception summary table. For each process, the table shows the number of subgroups that were analyzed and the number of subgroup summary statistics that fell outside the control limits of each chart (SAS Institute Inc. 2022):

The SPC Procedure

Means and Ranges Chart Summary for ALLPROCESSES						
			Means		Ranges	
processname	subgroupname	Number of Subgroups	Number > UCL	Number < LCL	Number > UCL	Number < LCL
Amount	Batch	23	0	0	0	0
Breakstrength	Sample	25	0	0	0	0
Delay	Day	11	3	1	2	2
Diameter	Batch	45	0	1	0	0
KWatts	Day	40	0	0	0	0
Partgap	Sample	41	0	0	0	0
Time	Lot	25	0	0	0	0
Weight	Lot	28	0	0	0	0

Figure 19 - PROC SPC Output

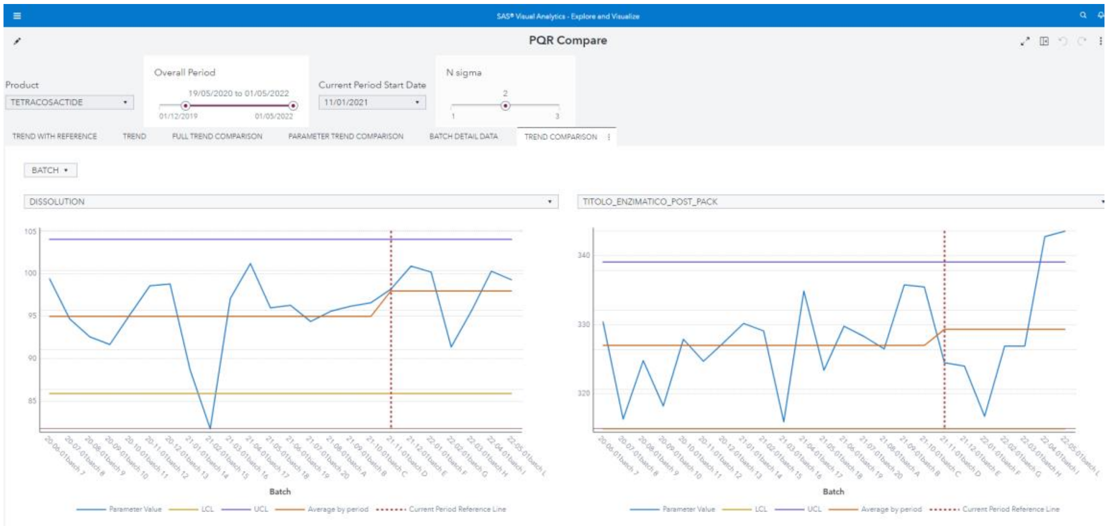
It is possible to process thousand on control charts in batch mode, mark the ones which are not in control or violates "western electric rules". By loading the corresponding data on SAS Visual Analytics we can plot the violating charts and analyze them in further detail with the flexible approach described above.

The approach lends itself to being used both:

- Tactically when it is useful to insert a control chart as an enrichment of a report focused on a broader topic.
- Strategically when it comes to adding an interactive display of exceptions to massive batch processes production of control charts.

In the second scenario, the goal is not to replace such procedures but to leverage all the feature and functions of SAS Visual Analytics about data visualization, report authoring, user profiled sharing and collaboration.

With a minimal training on SAS Visual Analytics and no coding, the analyst can produce his/her views, for example comparing the control chart of two parameters:



The possibilities are limitless and are typical of a self-service analytics tool.

CONCLUSION

In this paper we describe a real world requirement for a dynamic and flexible control chart to be used in the Product Quality Review Process of a Pharma company. We show how this can be easily implemented in SAS Visual Analytics. We provide motivations and benefits of this approach found in Healthcare literature. We then generalize the real word use case to the XS chart of SAS/QC proc Shewhart. We provide hints on how this approach can be applied to build the complete taxonomy of control charts. Lastly, we discuss how this approach can complement and enhance a SAS batch process production of control charts.

REFERENCES

Provost, L. P. and Murray, S. K. 2011. *The Health Care Data Guide: Learning from Data for Improvement*. San Francisco, CA: John Wiley & Sons

SAS Institute Inc. 2020. *SAS/QC® 15.2 User's Guide*. Cary, NC: SAS Institute Inc.

SAS Institute Inc. 2022. *SAS® Visual Statistics: Procedures*. Cary, NC: SAS Institute Inc

Jones, R. and Liu, L. 2018. "Leveraging SHEWHART Procedure Options to Monitor and Evaluate Improvements in Healthcare. " *Proceedings of the Midwest SAS User Group 2018 Conference*, Indianapolis, IN. Available at https://www.mwsug.org/proceedings/2018/HS/MWSUG-2018-HS-71.pdf.

Hessing, Ted. 2022. "Control Charts Study Guide." Accessed January 3, 2023. Available at https://sixsigmastudyguide.com/control-charts-study-guide.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Massimo Fabris
SAS Italy
massimo.fabris@sas.com

APPENDIX

```
/* How to calculate C4 and D2 values with SAS QC */

data control_chart_coefficients;
    do N = 2 to 40;
    C4 = c4(N);
    D2 = d2(N);
    output;
    end;
run;

/* Fixed table if SAS/QC is not licensed - many sources available */

/*
data control_chart_coefficients;
informat N best12. D2 best12. c4 best12.;
input N @;
input D2 @;
input C4 @;
datalines;
2 1.128 0.7979
3 1.693 0.8862
4 2.059 0.9213
5 2.326 0.9400
6 2.534 0.9515
7 2.704 0.9594
8 2.847 0.9650
9 2.970 0.9693
10 3.078 0.9727
;
run;

*/

data TurbineWithVaryingSubgroup;
informat Day date7.;
format Day YYMMDD10.;
label KWatts='KWatts';
input Day @;
do i=1 to 10;
input KWatts @;
output;
end;
drop i;
datalines;
04JUL94 3196 3507 4050 3215 3583 3617 3789 3180 3505 3454
05JUL94 3390 3562 3413 3193 3635 3179 3348 3199 3413 3562
05JUL94 3428 3320 3745 3426 3849 3256 3841 3575 3752 3347
06JUL94 3478 3465 3445 3383 3684 3304 3398 3578 3348 3369
06JUL94 3670 3614 3307 3595 3448 3304 3385 3499 3781 3711
07JUL94 3448 3045 3446 3620 3466 3533 3590 3070 3499 3457
07JUL94 3411 3350 3417 3629 3400 3381 3309 3608 3438 3567
08JUL94 3568 2968 3514 3465 3175 3358 3460 3851 3845 2983
08JUL94 3410 3274 3590 3527 3509 3284 3457 3729 3916 3633
09JUL94 3153 3408 3741 3203 3047 3580 3571 3579 3602 3335
```

```
09JUL94 3494 3662 3586 3628 3881 3443 3456 3593 3827 3573
10JUL94 3594 3711 3369 3341 3611 3496 3554 3400 3295 3002
11JUL94 3482 3546 3196 3379 3559 3235 3549 3445 3413 3859
11JUL94 3330 3465 3994 3362 3309 3781 3211 3550 3637 3626
12JUL94 3152 3269 3431 3438 3575 3476 3115 3146 3731 3171
12JUL94 3206 3140 3562 3592 3722 3421 3471 3621 3361 3370
13JUL94 3421 3381 4040 3467 3475 3285 3619 3325 3317 3472
13JUL94 2532 1835 3366 3492 3367 3619 1635 3263 3355 3510
14JUL94 3795 3872 3559 3432 3322 3587 3336 3732 3451 3215
14JUL94 3594 3410 3335 3216 3336 3638 3419 3515 3399 3709
15JUL94 3850 3431 3460 3623 3516 3810 3671 3602 3480 3388
15JUL94 3365 3845 3520 3708 3202 3365 3731 3840 3182 3677
16JUL94 3711 3648 3212 3664 3281 3371 3416 3636 3701 3385
16JUL94 3769 3586 3540 3703 3320 3323 3480 3750 3490 3395
17JUL94 3596 3436 3757 3288 1417 3331 3475 3600 3690 3534
18JUL94 3428 3760 3641 3393 3182 3381 3425 3467 3451 3189
18JUL94 3588 3484 3759 3292 3063 3442 3712 3061 3815 3339
19JUL94 3746 3426 3320 3819 3584 3877 3779 3506 3787 3676
19JUL94 3727 3366 3288 3684 3500 3501 3427 3508 3392 3814
20JUL94 3428 3760 3641 3393 3182 3381 3425 3467 3451 3189
20JUL94 3676 3475 3595 3122 3429 3474 3125 3307 3467 3832
20JUL94 3383 3114 3431 3693 3363 3486 3928 3753 3552 3524
22JUL94 3618 3324 3475 3621 3376 3540 3585 3320 3256 3443
23JUL94 3421 3787 3454 3699 3307 3917 3292 3310 3283 3536
23JUL94 3756 3145 3571 3331 3725 3605 3547 3421 3257 3574
;
run;

proc means data=TurbineDataWithVaryingSubgroup nway missing;
class Day;
var kwatts;
output out=statByday(drop=_type_ _freq_) n=n mean=mean std=std ;
run;

proc print;
run;


proc sql;
create table statBydayPlusCoeff
as
select a.*, b.C4
from
statByday as a
left join
control_chart_coefficients as b
on a.N = b.N
order by Day;
;
quit;

proc print;
run;
```

```
ods noproctitle;
ods graphics / imagemap=on;

proc shewhart data=TurbineWithVaryingSubgroup;
   xschart KWatts * Day / outtable= xsChartWithDiffgroups
outlimits=xsChartWithDiffgroupsLimits;
run;

proc print data=xsChartWithDiffgroups label;
var day _SUBN_ _SUBX_  _SUBS_ _MEAN_ _STDDEV_  _LCLX_ _UCLX_;
run;

/* calculate mean and Std Dev estimate with SAS Base */
proc sql;
create table WeightMean_StdDevEstimate
as
select sum(N *MEAN)/sum(N) as WeightMean, sum(STD / C4) / count(*) as
StdDevEstimate
from
statBydayPlusCoeff
;
quit;


/* Compare the results */

data compare;
   merge WeightMean_StdDevEstimate xsChartWithDiffgroupsLimits(keep=
_mean_ _stddev_);
   run;

data ChartWithDiffgroupsData;
   set xsChartWithDiffgroups;
   UCL = _mean_ + _SIGMAS_ * _STDDEV_ / (sqrt(_LIMITN_));
   LCL = _mean_ - _SIGMAS_ * _STDDEV_ / (sqrt(_LIMITN_));
   run;

proc print data=chartwithdiffgroupsdata;
var _UCLX_ UCL _LCLX_ LCL ;
run;
```