

Chapter 242

X-bar and R Charts

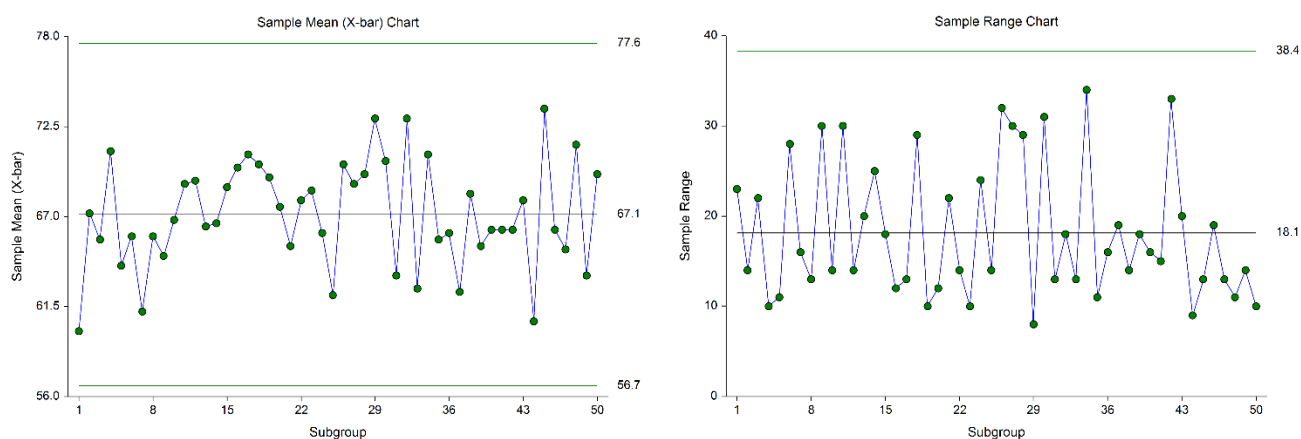
Introduction

This procedure generates X-bar and R control charts for variables. The format of the control charts is fully customizable. The data for the subgroups can be in a single column or in multiple columns. This procedure permits the defining of stages. For the X-bar chart, the center line can be entered directly or estimated from the data, or a sub-set of the data. Similarly, sigma may be estimated from the data or a standard sigma value may be entered. A list of out-of-control points can be produced in the output, if desired, and means and ranges may be stored to the spreadsheet.

X-bar and R Control Charts

X-bar and R charts are used to monitor the mean and variation of a process based on samples taken from the process at given times (hours, shifts, days, weeks, months, etc.). The measurements of the samples at a given time constitute a subgroup. Typically, an initial series of subgroups is used to estimate the mean and standard deviation of a process. The mean and standard deviation are then used to produce control limits for the mean and range of each subgroup. During this initial phase, the process should be in control. If points are out-of-control during the initial (estimation) phase, the assignable cause should be determined and the subgroup should be removed from estimation. Determining the process capability (see R & R Study and Capability Analysis procedures) may also be useful at this phase.

Once the control limits have been established of the X-bar and R charts, these limits may be used to monitor the mean and variation of the process going forward. When a point is outside these established control limits it indicates that the mean (or variation) of the process is out-of-control. An assignable cause is suspected whenever the control chart indicates an out-of-control process.



Other Control Charts for the Mean and Variation of a Process

Historically, the X-bar and R charts have been the most commonly used control charts for the process mean and process variation, in part because they are the simplest to calculate. A very similar pair of charts are the X-bar and s charts. The X-bar and s charts use the standard deviation of the subgroups rather than the range for estimating the standard deviation (σ). The X-bar and s charts are generally recommended over the X-bar and r charts when the subgroup sample size is moderately large ($n > 10$), or when the sample size is variable from subgroup to subgroup (Montgomery, 2013).

Two additional control charts available for monitoring the process mean are the cumulative sum (CUSUM) and exponentially weighted moving average (EWMA) charts. The CUSUM and EWMA charts differ from the X-bar charts in that they take into account the information of previous means at each point rather than just the current mean. The CUSUM and EWMA charts are more sensitive to smaller shifts in the mean since they use the cumulative information of the sequence of means. However, CUSUM and EWMA methods also assume a reliable estimate or known value for the true standard deviation is available.

When only a single response is available at each time point, then the individuals and moving range (I-MR) control charts can be used for early phase monitoring of the mean and variation. CUSUM and EWMA charts may also be used for single responses and are useful when small changes in the mean need to be detected.

Control Chart Formulas

Suppose we have k subgroups, each of size n . Let x_{ij} represent the measurement in the j^{th} sample of the i^{th} subgroup.

Formulas for the Points on the Chart

The i^{th} subgroup mean is calculated using

$$\bar{x}_i = \frac{\sum_{j=1}^n x_{ij}}{n},$$

and the subgroup range is calculated with

$$R_i = x_{i(n)} - x_{i(1)},$$

which is the smallest observation of the subgroup subtracted from the largest.

Estimating the X-bar Chart Center Line (Grand Mean)

In the X-bar and R Charts procedure, the grand average may be input directly, or it may be estimated from a series of subgroups. If it is estimated from the subgroups the formula for the grand average is

$$\bar{\bar{x}} = \frac{\sum_{i=1}^k \sum_{j=1}^{n_i} x_{ij}}{\sum_{i=1}^k n_i}.$$

If the subgroups are of equal size, the above equation for the grand mean reduces to

$$\bar{\bar{x}} = \frac{\sum_{i=1}^k \bar{x}_i}{k} = \frac{\bar{x}_1 + \bar{x}_2 + \cdots + \bar{x}_k}{k}.$$

Estimating Sigma

The true standard deviation (sigma) may be input directly, or it may be estimated from the ranges by

$$\hat{\sigma} = \frac{\bar{R}}{d_2}$$

where

$$\bar{R} = \frac{\sum_{i=1}^k R_i}{k}$$

$$d_2 = \frac{E(R)}{\sigma} = \frac{\mu_R}{\sigma}$$

The calculation of $E(R)$ requires the knowledge of the underlying distribution of the x_{ij} 's. Making the assumption that the x_{ij} 's follow the normal distribution with constant mean and variance, the values for d_2 are derived through the use of numerical integration. It is important to note that the normality assumption is used and that the accuracy of this estimate requires that this assumption be valid.

When n is one, we cannot calculate R_i since it requires at least two measurements. The procedure in this case is to use the ranges of successive pairs of observations. Hence, the range of the first and second observation is computed, the range of the second and third is computed, and so on. The average of these approximate ranges is used to estimate σ .

X-bar Chart Limits

The lower and upper control limits for the X-bar chart are calculated using the formulas

$$LCL = \bar{\bar{x}} - m \left(\frac{\hat{\sigma}}{\sqrt{n}} \right)$$

$$UCL = \bar{\bar{x}} + m \left(\frac{\hat{\sigma}}{\sqrt{n}} \right)$$

where m is a multiplier (usually set to 3) chosen to control the likelihood of false alarms (out-of-control signals when the process is in control).

Estimating the R Chart Center Line

If a standard sigma (standard deviation) value is entered by the user, the R Chart center line is computed using

$$\bar{R} = d_2 \sigma$$

If the standard deviation is estimated from a series of subgroups, the R chart center line is given by

$$\bar{R} = \frac{\sum_{i=1}^k R_i}{k}$$

R Chart Limits

The lower and upper control limits for the range chart are calculated using the formula

$$LCL = \bar{R} - md_3 \hat{\sigma}$$

$$UCL = \bar{R} + md_3 \hat{\sigma}$$

where m is a multiplier (usually set to 3) chosen to control the likelihood of false alarms, and d_3 is a constant (which depends on n) that is calculated by numerical integration and is based on the assumption of normality. The relation for d_3 is

$$d_3 = \frac{\sigma_R}{\sigma}$$

Runs Tests

The strength of control charts comes from their ability to detect sudden changes in a process that result from the presence of assignable causes. Unfortunately, the X-bar chart is poor at detecting drifts (gradual trends) or small shifts in the process. For example, there might be a positive trend in the last ten subgroups, but until a mean goes above the upper control limit, the chart gives no indication that a change has taken place in the process.

Runs tests can be used to check control charts for unnatural patterns that are most likely caused by assignable causes. Runs tests are sometimes called “pattern tests”, “out-of-control” tests, or “zones rules”.

While runs tests may be helpful in identifying patterns or smaller shifts in the mean, they also increase the likelihood of false positive indications. The rate of false positives is typically measured using the average run length (the average length of a run before a false positive is indicated). When several runs tests are used the average run length of the control chart becomes very short. Two alternatives to consider before using runs tests are the CUSUM and EWMA control charts. Runs tests are generally advised against when there is only one observation per subgroup. In this case, the rate of false positives is quite high (average run length is short).

In order to perform the runs tests, the control chart is divided into six equal zones (three on each side of the centerline). Since the control limit is three sigma limits (three standard deviations of the mean) in width, each zone is one sigma wide and is labeled A, B, or C, with the C zone being the closest to the centerline. There is a lower zone A and an upper zone A. The same is true for B and C. The runs tests look at the pattern in which points fall in these zones.

The runs tests used in this procedure are described below.

Test 1: Any Single Point Beyond Zone A

This runs test simply indicates a single point is beyond one of the two three-sigma limits.

Test 2: Two of Three Successive Points in Zone A or Beyond

This usually indicates a shift in the process average. Note that the two points have to be in the same Zone A, upper or lower. They cannot be on both sides of the centerline. The third point can be anywhere.

Test 3: Four of Five Successive Points in Zone B or Beyond

This usually indicates a shift in the process average. Note that the odd point can be anywhere.

Test 4: Eight Successive Points in Zone C or Beyond

All eight points must be on one side of the centerline. This is another indication of a shift in the process average.

Test 5: Fifteen Successive Points Fall in Zone C on Either Side of the Centerline

Although this pattern might make you think that the variation in your process has suddenly decreased, this is usually not the case. It is usually an indication of stratification in the sample. This happens when the samples come from two distinct distributions having different means. Perhaps there are two machines that are set differently. Try to isolate the two processes and check each one separately.

Test 6: Eight of Eight Successive Points Outside of Zone C

This usually indicates a mixture of processes. This can happen when two supposedly identical production lines feed a single production or assembly process. You must separate the processes to find and correct the assignable cause.

There are, of course, many other sets of runs tests that have been developed. You should watch your data for trends, zig-zags, and other nonrandom patterns. Any of these conditions could be an indication of an assignable cause and would warrant further investigation.

Issues in Using Control Charts

There are several additional considerations surrounding the use of control charts that will not be addressed here. Some important questions are presented below without discussion. For a full treatment of these issues, you should consider a statistical quality control text such as Ryan (2011) or Montgomery (2013).

Subgroup Size

How many items should be sampled for each subgroup? Some common values are 5, 10, and 20. How does the subgroup size affect my use of control charts? What about unequal subgroup sizes?

Dealing with Out-of-Control Points

How do you deal with out-of-control points once they have been detected? Should they be included or excluded in the process average and standard deviation?

Control Limit Multiplier

Three-sigma limits are very common. When should one consider a value other than three?

Startup Time

How many subgroups should be used to establish control for my process?

Normality Assumption

How important is the assumption of normality? How do I check this? Should I consider a transformation? (See also the Box-Cox Transformation and Capability Analysis procedures in **NCSS**.)

Data Structure

In this procedure, the data may be in either of two formats. The first data structure option is to have the data in several columns, with one subgroup per row.

Example Dataset

S1	S2	S3	S4	S5
2	6	3	8	5
8	8	7	7	9
6	2	2	4	3
5	6	7	6	10
48	2	6	5	0
.
.
.

The second data structure option uses one column for the response data, and either a subgroup size or a second column defining the subgroups.

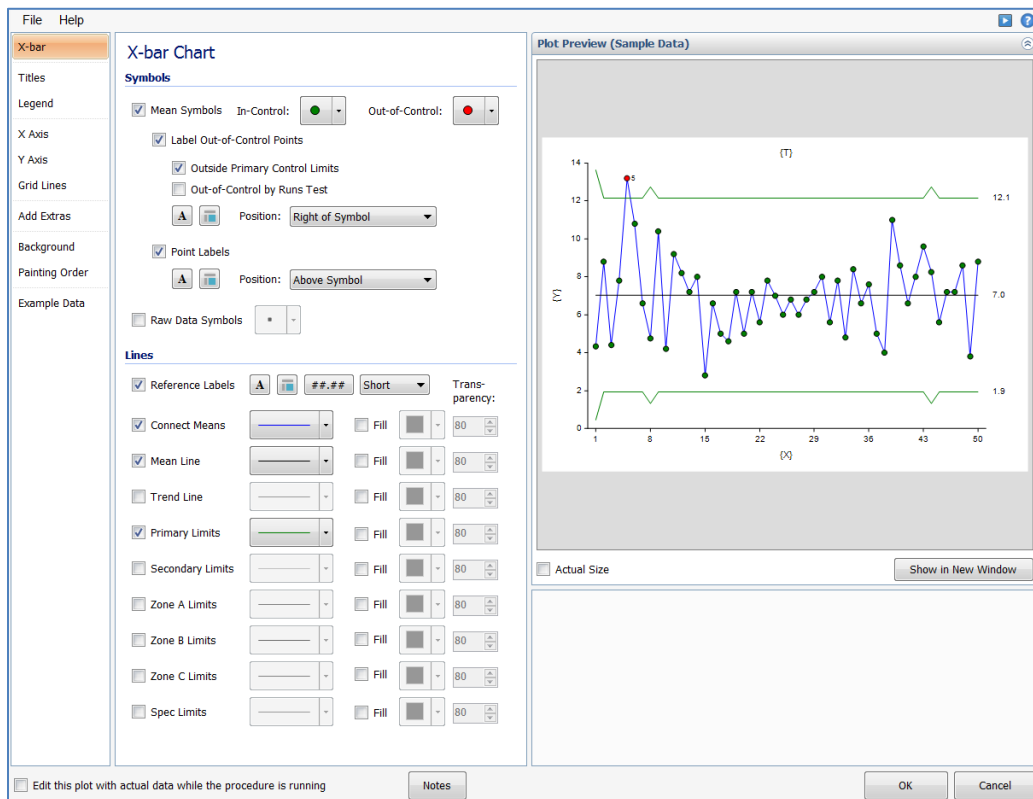
Alternative Example Dataset

Response	Subgroup
2	1
6	1
3	1
8	1
5	1
8	2
8	2
7	2
7	2
9	2
6	3
2	3
.	.
.	.
.	.

In the alternative example dataset, the Subgroup column is not needed if every subgroup is of size 5 and the user specifies 5 as the subgroup size. If there are missing values, the Subgroup column should be used, or the structure of the first example dataset.

Quality Control Chart Format Window Options

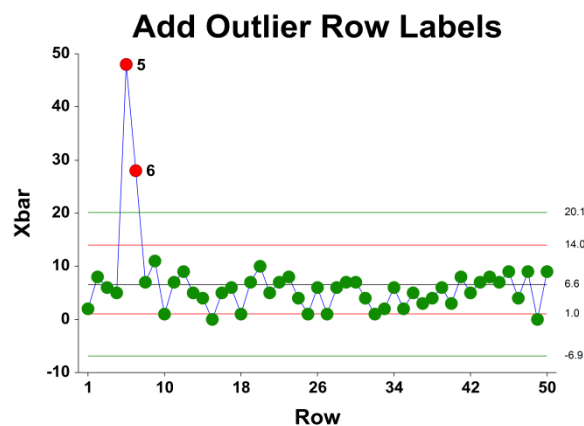
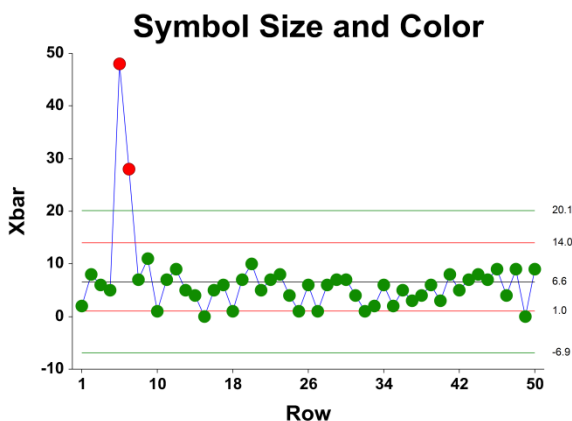
This section describes a few of the specific options available on the first tab of the control chart format window, which is displayed when a quality control chart format button is pressed. Common options, such as axes, labels, legends, and titles are documented in the Graphics Components chapter.



[Xbar] / [Range] Chart Tab

Symbols Section

You can modify the attributes of the symbols using the options in this section.



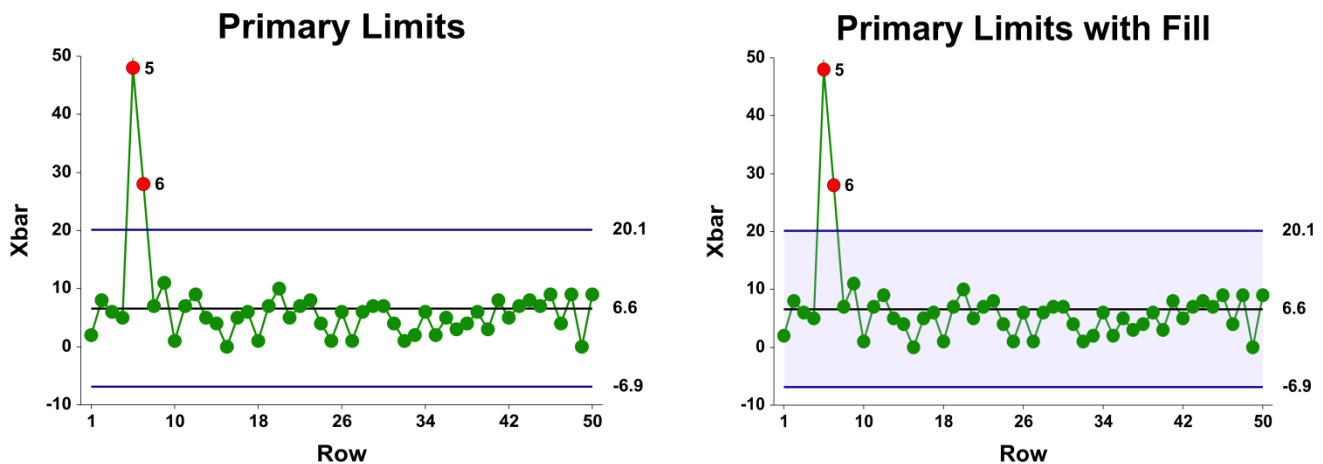
X-bar and R Charts

A wide variety of sizes, shapes, and colors are available for the symbols. The symbols for in-control and out-of-control points are specified independently. There are additional options to label out-of-control points. The label for points outside the primary control limits is the subgroup number. The label for points that are out-of-control based on the runs test is the number of the first runs test that is signaled by this point.

The user may also specify a column of point labels on the procedure variables tab, to be used to label all or some of the points of the chart. The raw data may also be shown, based on customizable raw data symbols.

Lines Section

You can specify the format of the various lines using the options in this section. Note that when shading is desired, the fill will be to the bottom for single lines (such as the mean line), and between the lines for pairs of lines (such as primary limits).



Lines for the zones, secondary limits, and specification limits are also specified here.

Titles, Legend, Numeric Axis, Group Axis, Grid Lines, and Background Tabs

Details on setting the options in these tabs are given in the Graphics Components chapter. The legend does not show by default but can easily be included by going to the Legend tab and clicking the Show Legend checkbox.

Example 1 – X-bar and R Charts Analysis (Phase I)

This section presents an example of how to run an initial X-bar and R Chart analysis to establish control limits. The data represent 50 subgroups of size 5. The data used are in the QC dataset. We will analyze the variables D1 through D5 of this dataset.

Setup

To run this example, complete the following steps:

1 Open the QC example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **QC** and click **OK**.

2 Specify the X-bar and R Charts procedure options

- Find and open the **X-bar and R Charts** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 1** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

```
Variables Tab
-----
Data Variables .....D1-D5
```

3 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Center Line Section

Center Line Section for Subgroups 1 to 50	
Number of Subgroups: 50	
Center Line	Estimate
Estimated Grand Average (X-bar-bar)	67.12
R-bar	18.14

This section displays the center line values that are to be used in the X-bar and R charts.

Estimated Grand Average (X-bar-bar)

This value is the average of all the observations. If all the subgroups are of the same size, it is also the average of all the X-bars.

R-bar

This is the average of the ranges.

Primary Control Limit Section

Primary Control Limit Section for Subgroups 1 to 50

These limits are based on a subgroup of size 5.

Chart Type	Primary Control Limits	
	Lower	Upper
X-bar	56.65682	77.58318
Range	0	38.35448

This report gives the lower and upper limits for each of the charts, corresponding to a specific subgroup size.

X-bar Lower and Upper Primary Control Limits

These limits are the primary control limits of the X-bar chart, as defined in the sub-section X-bar Chart Limits of the Control Chart Formulas section toward the beginning of this chapter.

Range Lower and Upper Primary Control Limits

These limits are the primary control limits of the R chart, as defined in the sub-section R Chart Limits of the Control Chart Formulas section toward the beginning of this chapter. Since the lower limit for the R chart is less than 0, it has been reset to 0.

Sigma Estimation Section

Sigma Estimation Section for Subgroups 1 to 50

Estimation Type	Estimated Value	Estimated Sigma
Ranges (R-bar)*	18.14	7.798796
Standard Deviations (s-bar)	7.365443	7.835698
Weighted Approach (s-bar)	7.902911	7.902911

* Indicates the estimation type used in this report.

This report gives the estimation of the population standard deviation (sigma) based on three estimation techniques. The estimation technique used for the plots in this procedure is based on the ranges.

Estimation Type

The formula for estimating sigma based on the ranges is shown earlier in this chapter in the Control Chart Formulas section. The formulas for the Standard Deviations method and Weighted Approach method are shown in the X-bar and s Charts chapter.

Estimated Value

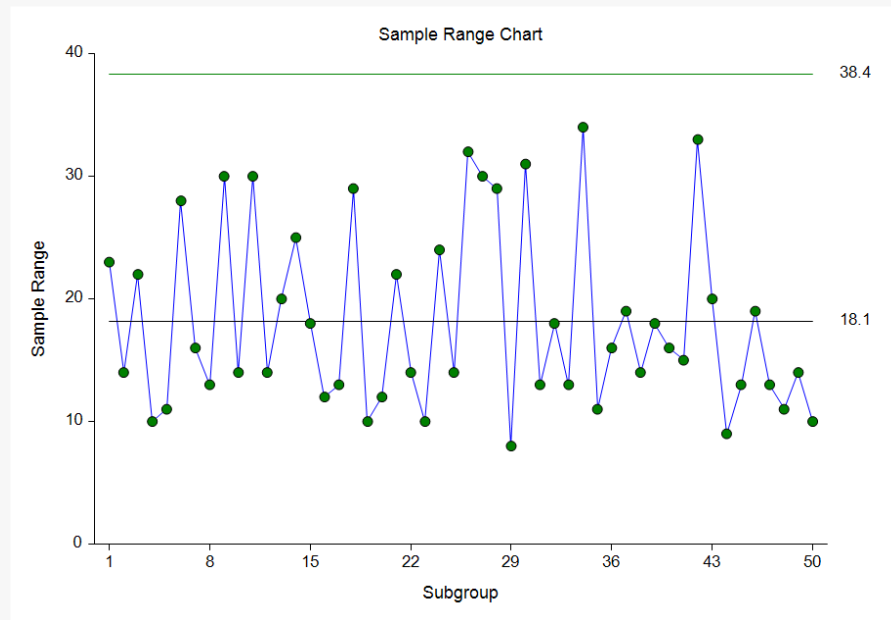
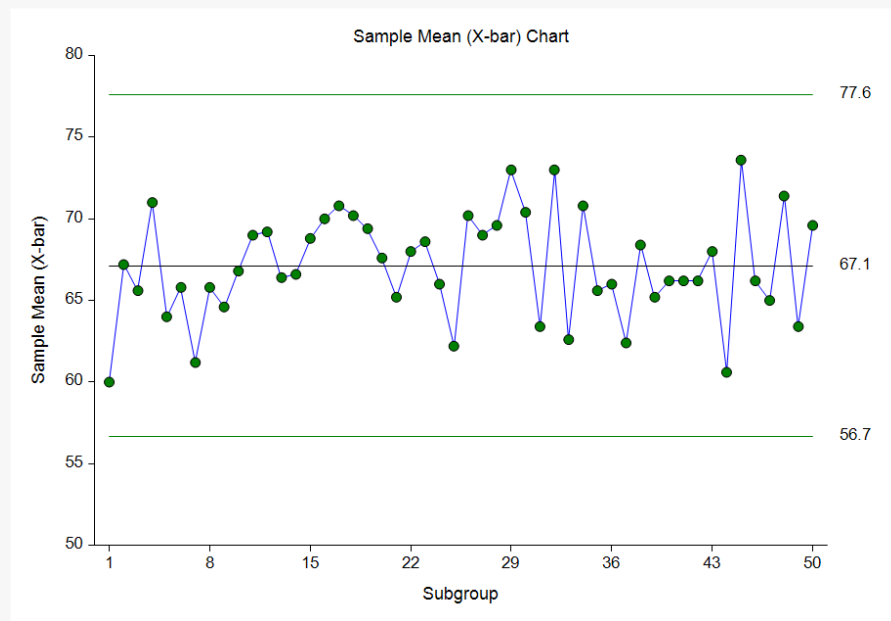
This column gives the R-bar and s-bar estimates based on the corresponding formulas.

Estimated Sigma

This column gives estimates of the population standard deviation (sigma) based on the corresponding estimation type.

X-bar and R Charts

Chart Section for Subgroups 1 to 50



The first plot shows the sample means, as well as the center line and control limits for the process mean, based on the 50 subgroups. This process appears to be in control. The second plot shows the ranges for each subgroup, as well as the corresponding center line and limits. The R chart seems to indicate the variation is also in control.

Out-of-Control List

Out-of-Control List for Subgroups 1 to 50

Subgroup	Mean	Range	Subgroup Label	Reason
30	70.4	31	30	Range: 4 of 5 in zone B or beyond

This report provides a list of the subgroups that failed one of the runs tests (including points outside the control limits). The report shows that subgroup 30 is the final point of 4 out of 5 points in Zone B. This run does not appear to indicate a clear shift in the process variation.

Example 2 – X-bar and R Charts Analysis (Phase II)

This section presents a continuation of the previous example. In this example the limits are based on the first 50 observations, but the means and ranges are monitored for an additional 100 subgroups. The data are given in the columns D1ext – D5ext of the QC dataset.

Setup

To run this example, complete the following steps:

1 Open the QC example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **QC** and click **OK**.

2 Specify the X-bar and R Charts procedure options

- Find and open the **X-bar and R Charts** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 2** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

```
Variables Tab
-----
Data Variables .....D1ext-D5ext
Specification Method .....First N rows (Enter N)
N .....50
```

3 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Center Line, Control Limits, and Estimation Sections

Center Line Section for Subgroups 1 to 50	
Number of Subgroups: 50	
Center Line	Estimate
Estimated Grand Average (X-bar-bar)	67.12
R-bar	18.14

X-bar and R Charts

Primary Control Limit Section for Subgroups 1 to 50

These limits are based on a subgroup of size 5.

Primary Control Limits		
Chart Type	Lower	Upper
X-bar	56.65682	77.58318
Range	0	38.35448

Sigma Estimation Section for Subgroups 1 to 50

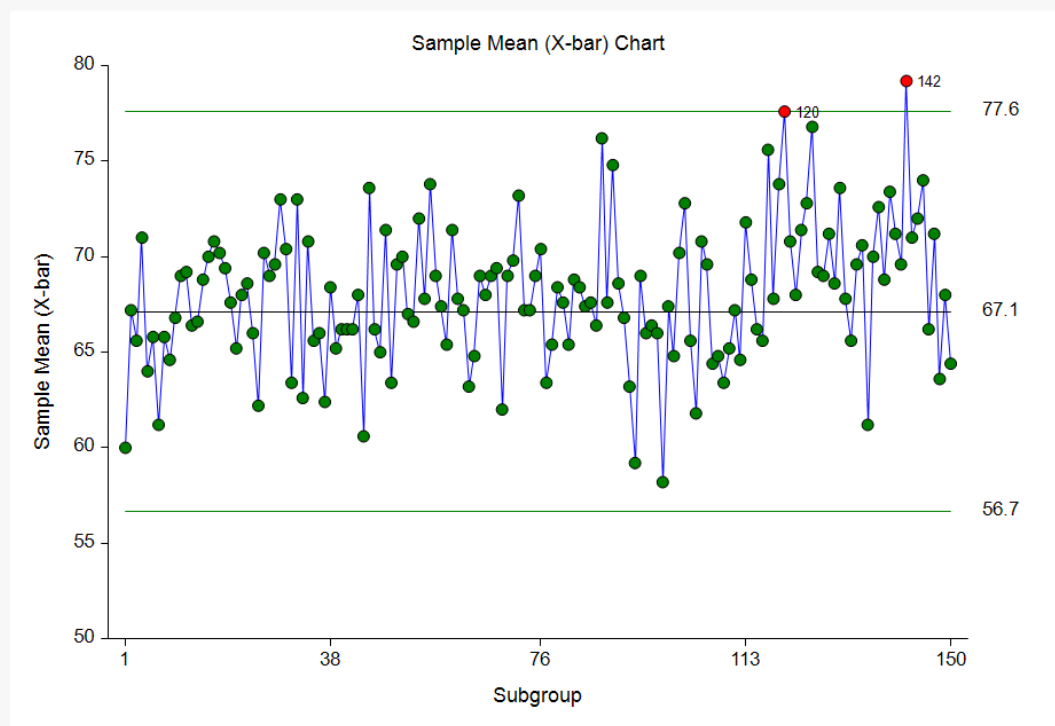
Estimation Type	Estimated Value	Estimated Sigma
Ranges (R-bar)*	18.14	7.798796
Standard Deviations (s-bar)	7.365443	7.835698
Weighted Approach (s-bar)	7.902911	7.902911

* Indicates the estimation type used in this report.

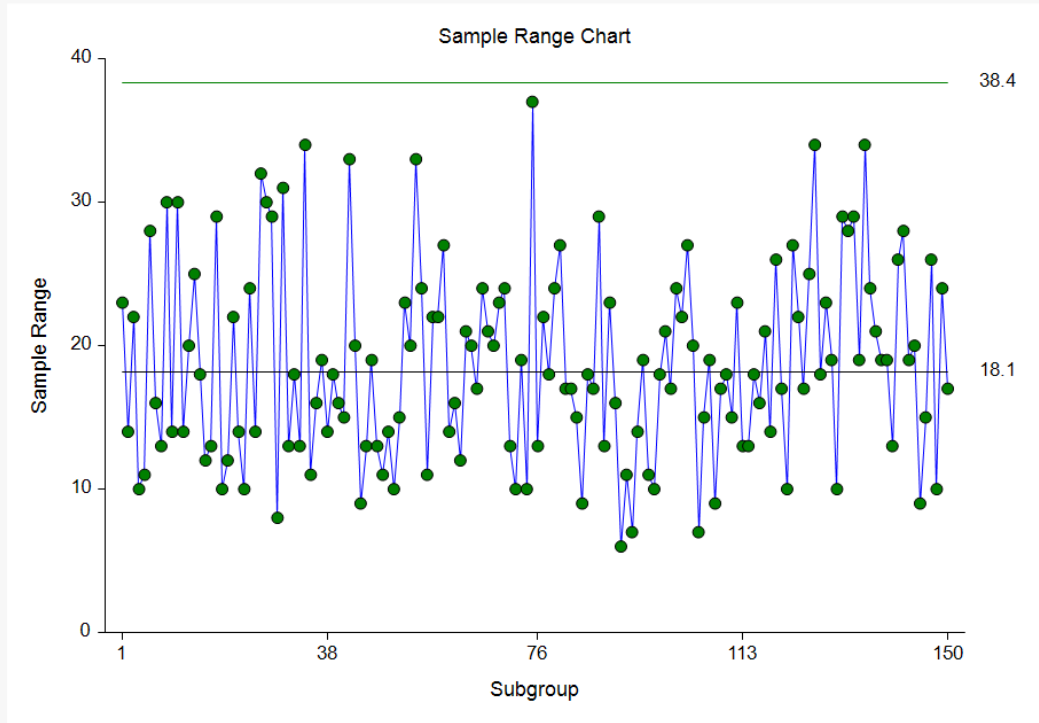
Since the first 50 subgroups are the same as those of Example 1, and since only the first 50 subgroups are used in the calculations for this run, the results for these sections are the same as those of Example 1.

X-bar and R Charts

Chart Section for Subgroups 1 to 150



X-bar and R Charts



These plots have the same limits as those of Example 1. It appears there may have been a shift in process for the last 30 or 40 subgroups, as evidenced by the large majority of points above the center line and out-of-control signal from subgroups 120 and 142.

Out-of-Control List

Out-of-Control List for Subgroups 1 to 150

Subgroup	Mean	Range	Subgroup Label	Reason
30	70.4	31	30	Range: 4 of 5 in zone B or beyond
89	74.8	23	89	X-bar: 2 of 3 in zone A
120	77.6	17	120	X-bar: beyond control limits
121	70.8	10	121	X-bar: 4 of 5 in zone B or beyond
123	71.4	22	123	X-bar: 4 of 5 in zone B or beyond
124	72.8	17	124	X-bar: 4 of 5 in zone B or beyond
125	76.8	25	125	X-bar: 4 of 5 in zone B or beyond
126	69.2	34	126	X-bar: 8 in zone C or beyond
127	69	18	127	X-bar: 8 in zone C or beyond
128	71.2	23	128	X-bar: 8 in zone C or beyond
129	68.6	19	129	X-bar: 8 in zone C or beyond
130	73.6	10	130	X-bar: 8 in zone C or beyond
131	67.8	29	131	X-bar: 8 in zone C or beyond
135	61.2	34	135	Range: 4 of 5 in zone B or beyond
138	68.8	19	138	Range: 8 in zone C or beyond
.
.
.

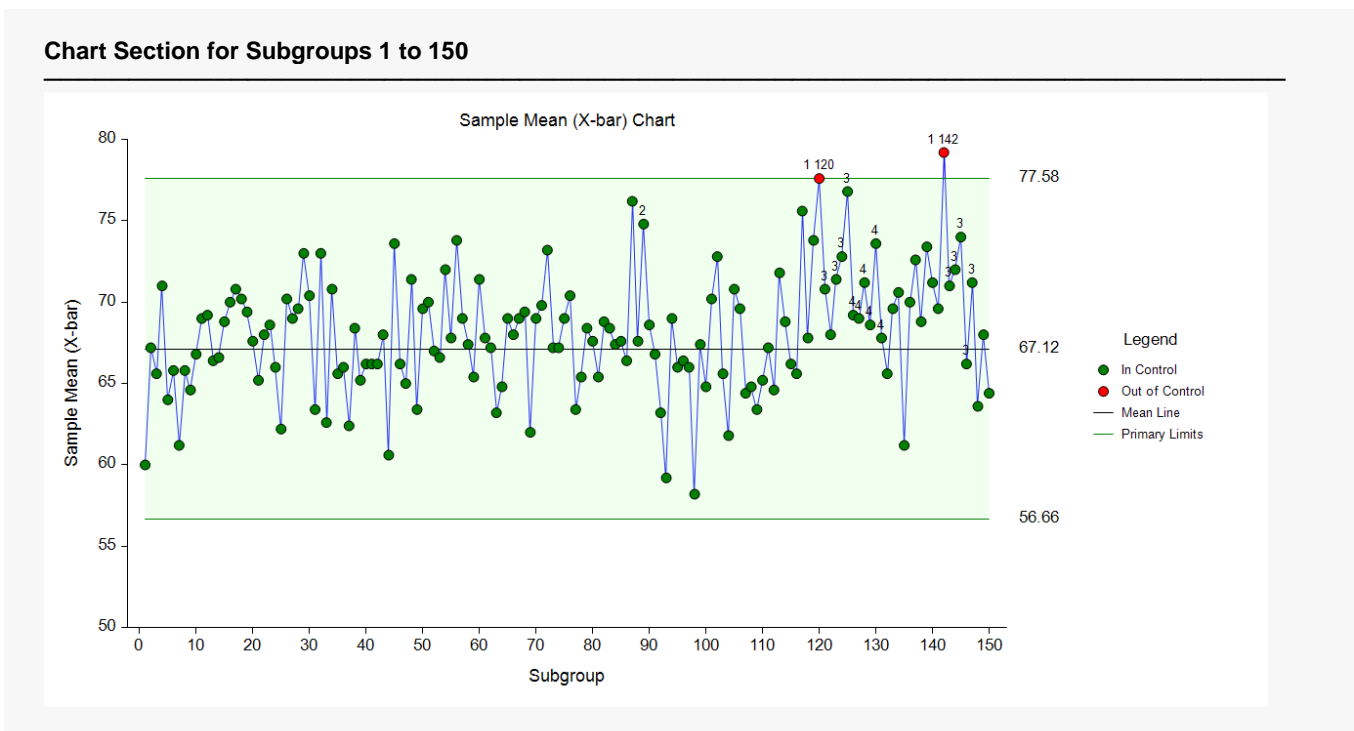
This list indicates a large number of out-of-control signals (by Runs tests) for subgroups 120 and beyond.

Example 3 – X-bar Chart with Additional Formatting

This example uses the same setup as Example 2, except that a variety of improvements are made in the plot format. These improvements are made by clicking the X-bar Chart format button on the X-bar & R charts tab.

The settings for this example are stored in the **Example 3** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

X-bar Chart



As shown here, a variety of enhancements can be made to the formatting of the control charts to make the chart as easy to read as possible. The numbers above the points near the end represent the number of the first runs test that is signaled by that point.

Example 4 – X-bar and R Charts with Stages

This section presents X-bar and R charts for the scenario of multiple stages in a process. In the QC Stage dataset a significant change had been made in the process after subgroup 70. This change merits a new evaluation of the process mean and standard deviation, beginning at subgroup 71. In this example, the process mean and standard deviation for each stage are to be determined based on the first 30 subgroups.

Setup

To run this example, complete the following steps:

1 Open the QC Stage example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **QC Stage** and click **OK**.

2 Specify the X-bar and R Charts procedure options

- Find and open the **X-bar and R Charts** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 4** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Variables Tab

Data Variables **D1-D5**
 Number of Stages..... **Multiple Stages**
 Stage Specification **Stage Variable**
 Stage Variable **Stage**
 Specification Method **Keep Rows Variable**
 Use all rows where variable..... **CalcRows**
 equals **1**

3 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Center Line, Control Limits, and Estimation Sections

Center Line Section for Subgroups 1 to 70 where CalcRows equals 1

Number of Subgroups: 30

Center Line	Estimate
Estimated Grand Average (X-bar-bar)	51.29908
R-bar	14.70471

X-bar and R Charts

Primary Control Limit Section for Subgroups 1 to 70 where CalcRows equals 1

These limits are based on a subgroup of size 5.

Chart Type	Primary Control Limits	
	Lower	Upper
X-bar	42.81738	59.78078
Range	0	31.09104

Sigma Estimation Section for Subgroups 1 to 70 where CalcRows equals 1

Estimation Type	Estimated Value	Estimated Sigma
Ranges (R-bar)*	14.70471	6.321887
Standard Deviations (s-bar)	5.921397	6.299455
Weighted Approach (s-bar)	6.408969	6.408969

* Indicates the estimation type used in this report.

Center Line Section for Subgroups 71 to 120 where CalcRows equals 1

Number of Subgroups: 30

Center Line	Estimate
Estimated Grand Average (X-bar-bar)	46.23898
R-bar	10.6469

Primary Control Limit Section for Subgroups 71 to 120 where CalcRows equals 1

These limits are based on a subgroup of size 5.

Chart Type	Primary Control Limits	
	Lower	Upper
X-bar	40.09783	52.38013
Range	0	22.51137

Sigma Estimation Section for Subgroups 71 to 120 where CalcRows equals 1

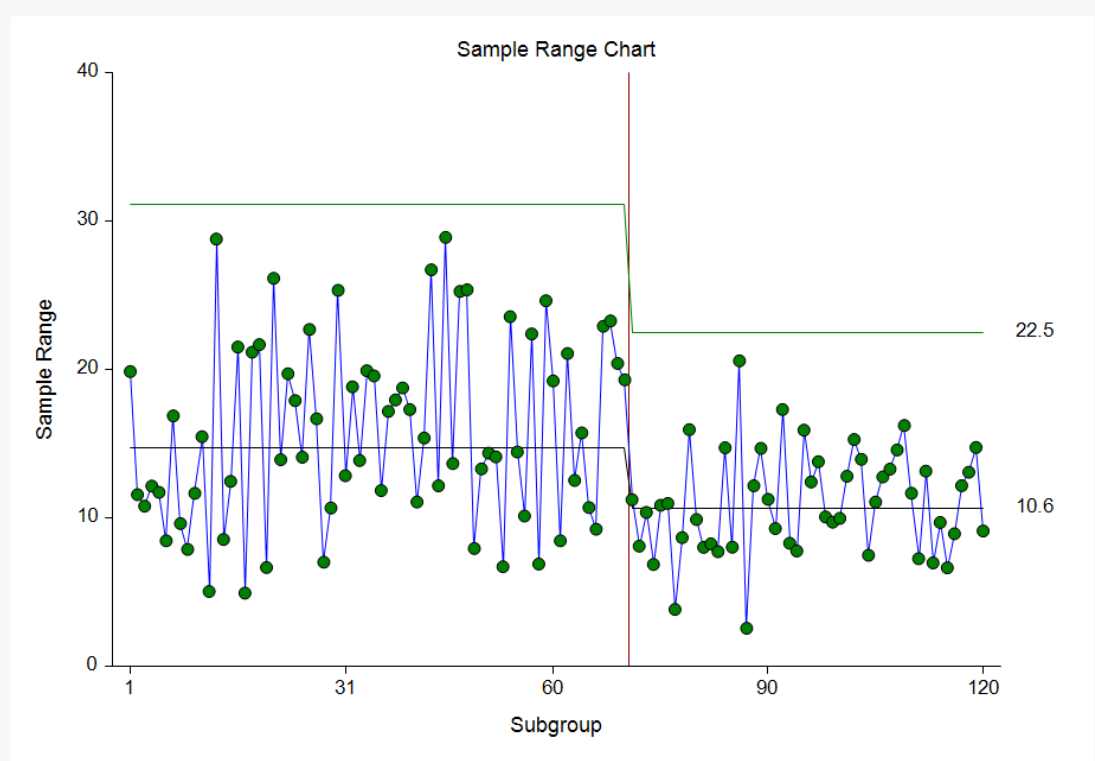
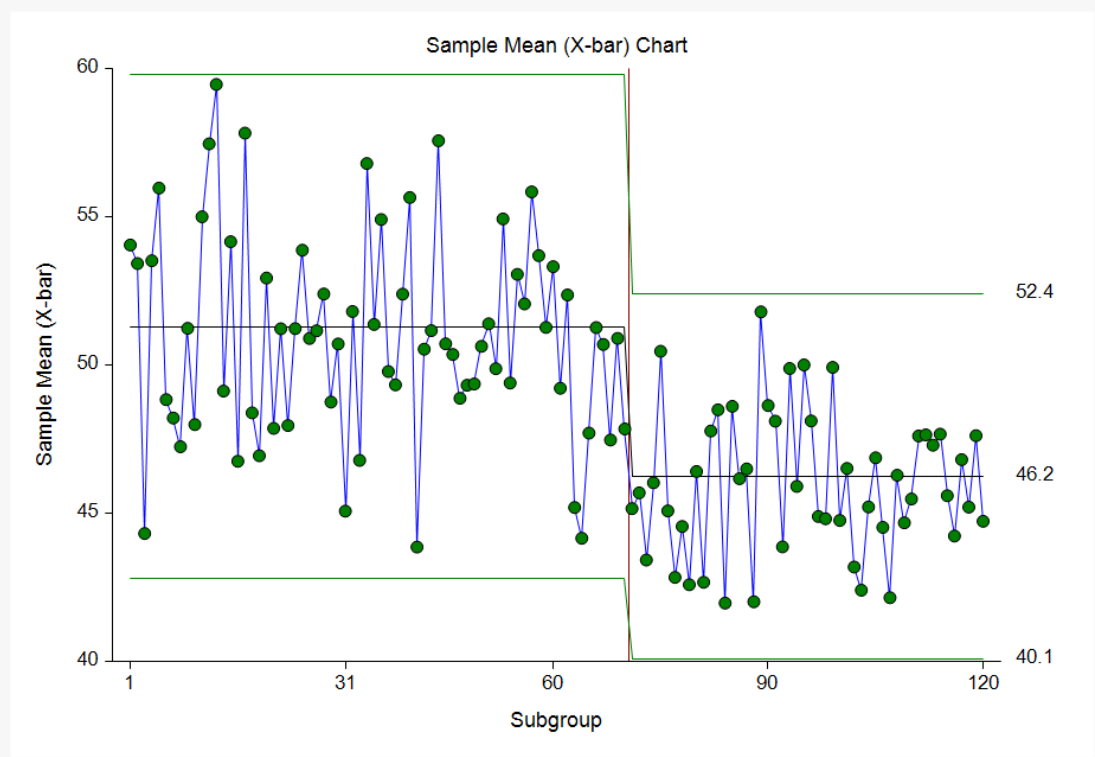
Estimation Type	Estimated Value	Estimated Sigma
Ranges (R-bar)*	10.6469	4.577343
Standard Deviations (s-bar)	4.244057	4.515023
Weighted Approach (s-bar)	4.490183	4.490183

* Indicates the estimation type used in this report.

This report shows the center line, limit, and sigma estimates for each of the two stages. The center line estimate changes from about 51 to about 46. There also appears to be a change in sigma of about 6.3 to about 4.5.

X-bar and R Charts

Chart Section for Subgroups 1 to 120



The change in process mean and variation from the first stage to the second is reflected in the X-bar and R charts.