CONSEQUENCES OF OUTLIERS IN VARIABLE CONTROL CHARTS

Weeraratne N.C.
Department of Economics & Statistics
SUSL, BelihulOya, Sri Lanka

ABSTRACT

The Control Chart is a statistical tool to determine if a process is in control or a graph used to determine how the process changes over time. There are two types of Control Charts. Among them, Variable Control Chart is the most famous type in Statistical Quality Control (SQC). Variable Control Charts deal with items that can be measured. For instance, $\bar{X}$ *and* R *Control Chart &* $\bar{X}$ *and* S *Control Chart are the most popular Variable Control Charts. This paper shows the issues of* $\bar{X}$ *and* R *Control Chart &* $\bar{X}$ *and* S *Control Chart. The* $\bar{X}$ *and* R *Control Chart &* $\bar{X}$ *and* S *Control Chart can sometimes be misleading when there are extremely high or low values.*

Key Words: *$\bar{X}$ and R Control Chart, $\bar{X}$ and S Control Chart, Outliers, SQC*

---

1. INTRODUCTION

Statistical Quality Control (SQC) refers to use statistical techniques for measuring and improving the quality of processes. It includes Statistical Process Control (SPC) in addition to other techniques, such as sampling plans, experimental design, variation reduction, process capability analysis and process improvement plans. In order to assess the quality of a product we need to make measurement(s) on the product and monitor their change over time [1]. The important measures of the quality of a product are called its quality characteristics. The quality characteristic may be a continuous measurement such as length, weight, time … etc. or it could be a discrete measurement such as number of errors, successes/ failures … etc. Variation exists in all processes and it is not possible to eliminate 100%, but it can be minimized at a greater extend. Generally there are two causes of variation. Chance or Common Cause Variability is a variation that is random in nature. Chance or natural variability is a cumulative effect of many small, essentially uncontrollable causes. This variation cannot be completely eliminated. For instance; slight differences in process variables like length, weight, service time. Assignable Cause Variability is a other kinds of variability could be present in the output of a process. It can be identified and eliminated. The variation due to assignable causes may be due to; difference among machines, difference among operators, difference among materials and poor employee training [4,5,6]. A process that is operating with only random or chance causes of variation present is said to be in statistical control process. A process that is operating in the presence of assignable causes is said to be statistical out of control process. Control Chart was constructed from historical data, the purpose of control chart is to help distinguish between natural variations and variations due to assignable causes [2]. Variable Control Chart and Attribute Control are the major types of Control Charts in Statistical Quality Control. For quality characteristics that have continuous dimensions, we develop a Variable Control Chart and for quality characteristics that have discrete dimensions, we develop an Attribute Control Chart. Usually, In Variable Control Charts, charts for control the dispersion of the process and charts for control the central tendency of the process must be used together. The theory of control chart was introduced by Walter A. Shewart and the general model for control charts as follows [3];

$$UCL = \mu_w + L\sigma_w$$
$$CL = \mu_w$$
$$LCL = \mu_w - L\sigma_w$$

Where;
$w = Sample\ statistic\ that\ measures\ some\ quality\ characteristics\ of\ interest$
$L = The\ distance\ of\ the\ control\ limit\ from\ the\ center\ line\ express\ in\ standard\ deviation\ units$
$\mu_w = Mean\ of\ w\ and\ \sigma_w = Standard\ deviation\ of\ w$

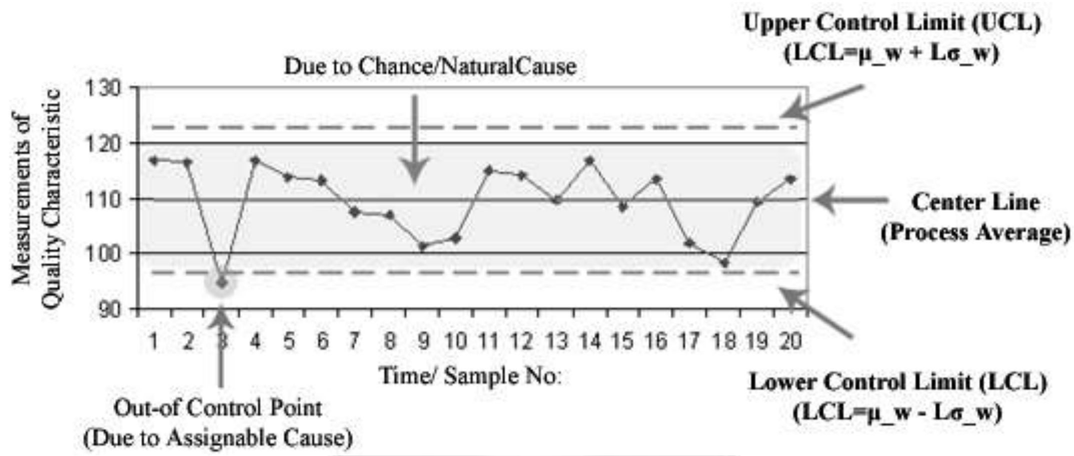


Figure 1: General Model for Control Chart

$\bar{X}$ and R Contrl Chart & $\bar{X}$ and S Contrl Chart (constant sample size)

$\bar{X}$ Charts are to control the central tendency of the process and R and S Charts are to control the dispersion of the process. These two control charts must be used together. $\bar{X}$ Chart detects shift in central tendency, but $\bar{X}$ Chart does not detect the increase in dispersion. R and S Chart does not detect shift in central tendency, but R and S Chart detects the increase in dispersion. Control limits based on sample values as follows;

Control limits for the $\bar{X}$ Chart (R):

$$UCL = \bar{\bar{X}} + A_2\bar{R}$$
$$CL = \bar{\bar{X}}$$
$$LCL = \bar{\bar{X}} - A_2\bar{R}$$

Control limits for the R Chart:

$$UCL = D_4\bar{R}$$
$$CL = \bar{R}$$
$$LCL = D_3\bar{R}$$

Control limits for the $\bar{X}$ Chart (S):

$$UCL = \bar{\bar{X}} + A_3\bar{S}$$
$$CL = \bar{\bar{X}}$$
$$LCL = \bar{\bar{X}} - A_3\bar{S}$$

Control limits for the S Chart:

$$UCL = B_4\bar{S}$$
$$CL = S$$
$$LCL = B_3\bar{S}$$

Where;

$$\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n}\ ; n = sample\ size\ and\ \bar{\bar{X}} = \frac{\sum_{i=1}^{m} \bar{x}_i}{m}\ ; m = no{:}of\ samples$$

$$R = x_{max} - x_{min}\ and\ \bar{R} = \frac{\sum_{i=1}^{m} R_i}{m}$$

$$S = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}\ \ and\ \bar{S} = \frac{\sum_{i=1}^{m} \bar{s}_i}{m}$$

$$A_2 = \frac{3}{d_2\sqrt{n}}\ ;\ d_2 = mean\ of\ relative\ range\ (w = R/\sigma)$$

$$B_3 = 1 - \frac{3}{c_4\sqrt{2(n-1)}}\ \ and\ \ B_4 = 1 + \frac{3}{c_4\sqrt{2(n-1)}}$$

$$D_3 = \left(1 - \frac{3d_3}{d_2}\right)\ and\ D_4 = \left(1 + \frac{3d_3}{d_2}\right)\ ;\ d_3 = standard\ deviation\ of\ relative\ range\ (w = R/\sigma)$$

In interpreting behavior on the $\bar{X}$ control chart, we must first determine whether or not the R/S control chart is statistically in control. Usually don't endeavor to interpret the $\bar{X}$ control chart when the R control chart indicates a statistically out of control condition. If both the $\bar{X}$ control chart and R/S control charts present a pattern, the best way is to eliminate the R/S control chart assignable causes first. In many cases, this will automatically ignore the pattern on the $\bar{X}$ control chart. [3]

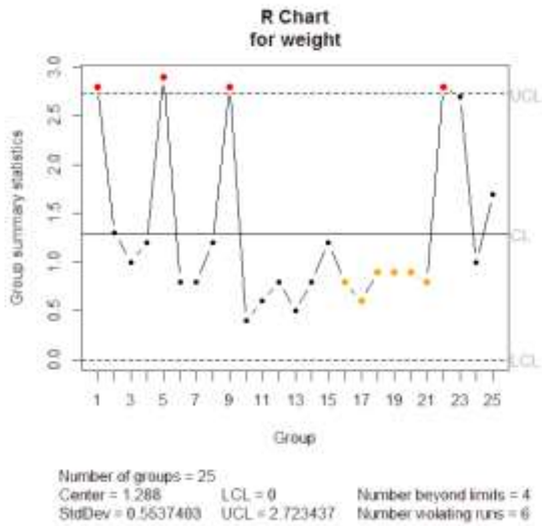
2. RESULTS AND FINDINGS

Illustration 01: The net weight (in grams) of a yam powder product is to be monitored by $\bar{X}$ control chart and R control chart and $\bar{X}$ control chart and S control chart using a sample size of n=5. Data for 25 preliminary samples and particular quality control charts are shown in following outputs.

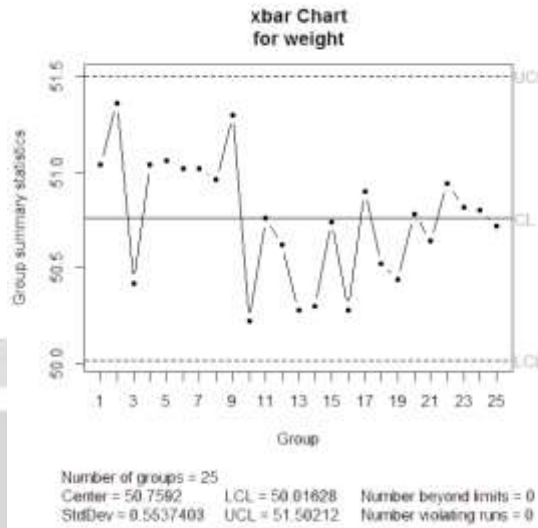
Table 1: Net weight (in grams) of a yam powder product – Illustration 01

X1	X2	X3	X4	X5	$\bar{X}_i$	R_i	S_i
51.8	50	50.1	52.8	50.5	51.04	**2.8**	**1.217785**
51.2	50.8	51.3	51.4	52.1	51.36	1.3	0.472229
50.1	50.5	50.5	50	51	50.42	1	0.396232
50.3	50.8	51.5	51.4	51.2	51.04	1.2	0.49295
50	50.2	51.3	52.9	50.9	51.06	**2.9**	**1.154556**
51.1	51.2	50.8	51.4	50.6	51.02	0.8	0.319374
50.8	50.9	50.7	51.2	51.5	51.02	0.8	0.327109
50.2	51.2	50.8	51.4	51.2	50.96	1.2	0.477493
50.9	50	51	52.8	51.8	51.3	**2.8**	1.053565
50.2	50.3	50.2	50	50.4	50.22	0.4	0.148324
50.7	50.8	50.9	50.4	51	50.76	0.6	0.230217
50.7	50.8	50.4	50.2	51	50.62	0.8	0.319374
50.2	50.3	50.5	50.4	50	50.28	0.5	0.192354
50.3	50.2	50.8	50.2	50	50.3	0.8	0.3
50.6	50.4	50.2	51.1	51.4	50.74	1.2	0.497996
50	50	50.4	50.8	50.2	50.28	0.8	0.334664
51.1	51.2	50.9	50.6	50.7	50.9	0.6	0.254951
50.7	50.8	50.9	50	50.2	50.52	0.9	0.396232
50.8	50.2	50	50.3	50.9	50.44	0.9	0.391152
51	51.2	50.8	50.6	50.3	50.78	0.9	0.349285
50.2	50.9	50.5	51	50.6	50.64	0.8	0.320936
50.1	50.2	50.8	50.7	52.9	50.94	**2.8**	**1.137102**
52.7	50.2	50.3	50	50.9	50.82	2.7	**1.103177**
50.8	50.7	50.2	51.2	51.1	50.8	1	0.3937
50.3	50.2	50.1	51.2	51.8	50.72	1.7	0.746324

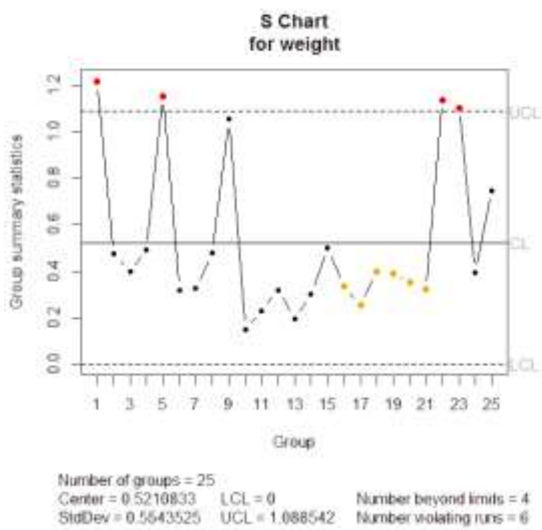
```
> q <- qcc(weight, type="R", nsigmas=3)
```




```
> q <- qcc(weight, type="xbar", nsigmas=3)
```




```
> q <- qcc(weight, type="S", nsigmas=3)
```




```
> q <- qcc(weight, type="xbar", nsigmas=3)
```

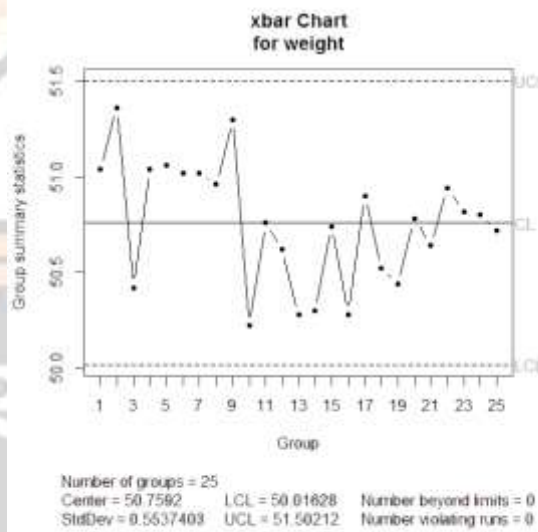


Figure 2: $\bar{X}$ *and R Control Chart (Illustration − 01)*

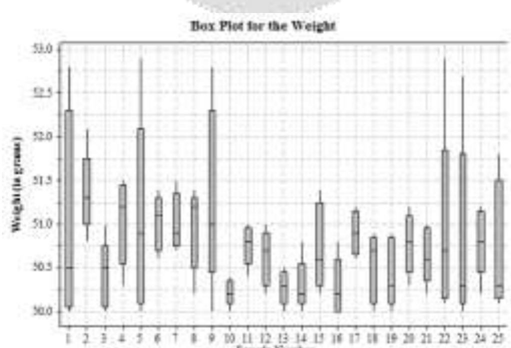


Figure 3: *Box plot of Weight (Illustration − 01)*

When the 25 sample ranges are plotted on the R control chart in Fig. 2 there is an indication of an out-of control condition. Since the R and S control chart indicates that process variability is out-of control, we may not go to discuss about the $\bar{X}$ control chart. Ultimately we can say this process is statistically out-of control process.

Illustration 02: The net weight (in grams) of a yam powder product is to be monitored by $\bar{X}$ control chart and R control chart and $\bar{X}$ control chart and S control chart using a sample size of n=5. Data for 25 preliminary samples and particular quality control charts are shown in following outputs (Note: Sample No: 22; X5 is a outlier value).

Table 2: Net weight (in grams) of a yam powder product – Illustration 02

X1	X2	X3	X4	X5	$\bar{X}_i$	R_i	S_i
51.8	50	50.1	52.8	50.5	51.04	2.8	1.217785
51.2	50.8	51.3	51.4	52.1	51.36	1.3	0.472229
50.1	50.5	50.5	50	51	50.42	1	0.396232
50.3	50.8	51.5	51.4	51.2	51.04	1.2	0.49295
50	50.2	51.3	52.9	50.9	51.06	2.9	1.154556
51.1	51.2	50.8	51.4	50.6	51.02	0.8	0.319374
50.8	50.9	50.7	51.2	51.5	51.02	0.8	0.327109
50.2	51.2	50.8	51.4	51.2	50.96	1.2	0.477493
50.9	50	51	52.8	51.8	51.3	2.8	1.053565
50.2	50.3	50.2	50	50.4	50.22	0.4	0.148324
50.7	50.8	50.9	50.4	51	50.76	0.6	0.230217
50.7	50.8	50.4	50.2	51	50.62	0.8	0.319374
50.2	50.3	50.5	50.4	50	50.28	0.5	0.192354
50.3	50.2	50.8	50.2	50	50.3	0.8	0.3
50.6	50.4	50.2	51.1	51.4	50.74	1.2	0.497996
50	50	50.4	50.8	50.2	50.28	0.8	0.334664
51.1	51.2	50.9	50.6	50.7	50.9	0.6	0.254951
50.7	50.8	50.9	50	50.2	50.52	0.9	0.396232
50.8	50.2	50	50.3	50.9	50.44	0.9	0.391152
51	51.2	50.8	50.6	50.3	50.78	0.9	0.349285
50.2	50.9	50.5	51	50.6	50.64	0.8	0.320936
50.1	50.2	50.8	50.7	**70** *	54.36	**19.9**	**8.748314**
52.7	50.2	50.3	50	50.9	50.82	2.7	1.103177
50.8	50.7	50.2	51.2	51.1	50.8	1	0.3937
50.3	50.2	50.1	51.2	51.8	50.72	1.7	0.746324

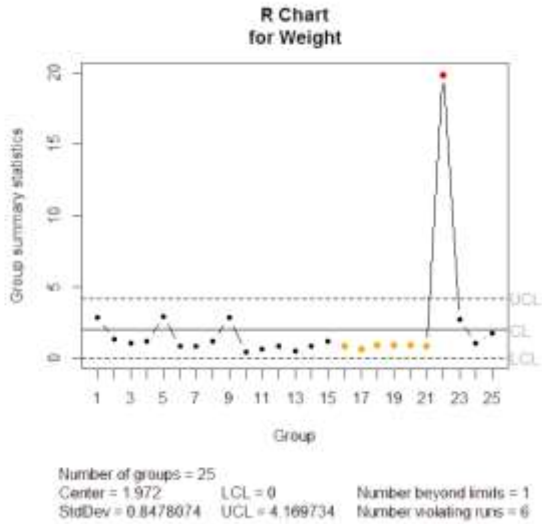
* Extremely Large Net Weight

$$\bar{\bar{X}} = \frac{\sum_{i=1}^{m} \bar{x}_i}{m} \qquad \bar{R} = \frac{\sum_{i=1}^{m} R_i}{m} \qquad \bar{S} = \frac{\sum_{i=1}^{m} \bar{s}_i}{m}$$

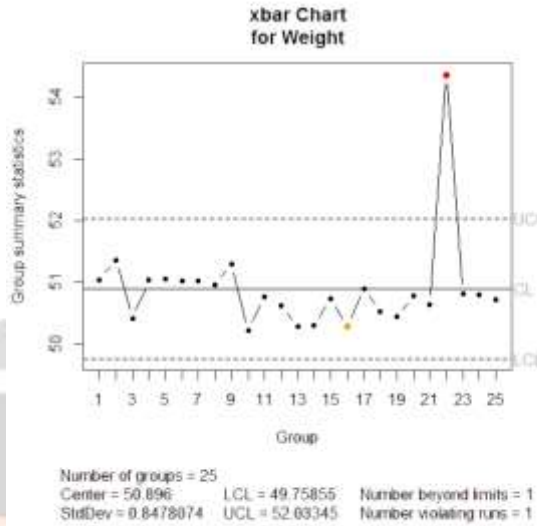
$\bar{\bar{X}} = 50.896$ $\bar{R} = 1.972$ $\bar{S} = 0.82553$

Due to the above net weight (in grams) of a yam powder product can have outliers (Sample No: 22) that are widely off the other net weight. So, the calculated descriptive statistics might not give a true indication of the behavior of net weight data.

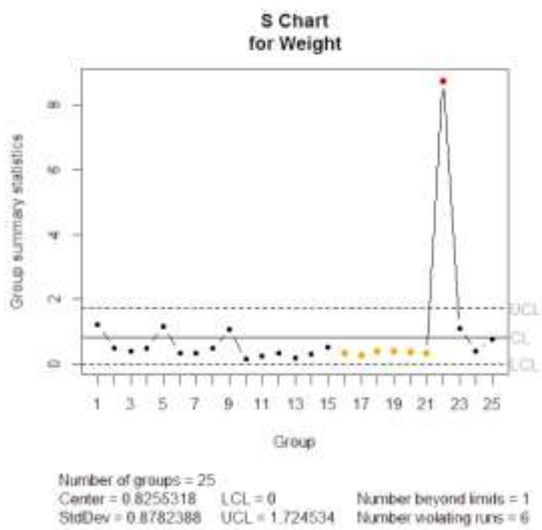
```
> q <- qcc(weight, type="R", nsigmas=3)
```




```
> q <- qcc(weight, type="xbar", nsigmas=3)
```




```
> q <- qcc(weight, type="S", nsigmas=3)
```




```
> q <- qcc(weight, type="xbar", nsigmas=3)
```

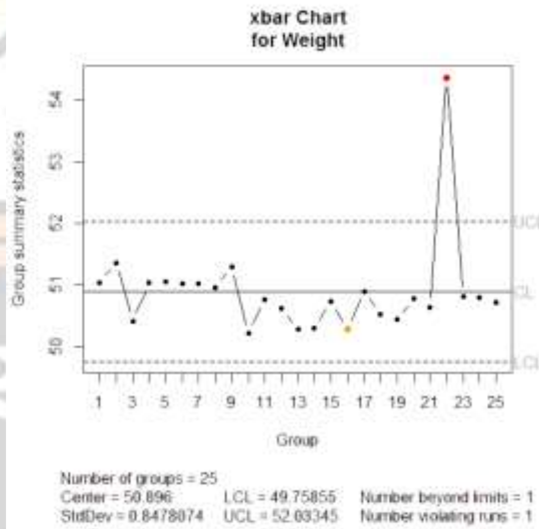


Figure 4: $\bar{X}$ *and R/S Control Chart (Illustration − 02)*

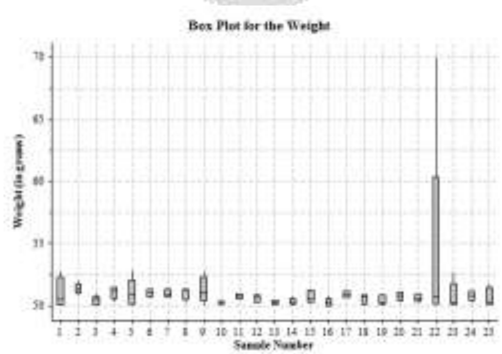


Figure 5: *Box plot of Weight (Illustration − 01)*

When the 25 sample ranges are plotted on the R control chart in Fig. 4 there is an indication of an out-of control condition. Since the R and S control chart indicates that process variability is out-of control, we may not go to discuss about the $\bar{X}$ control chart. Ultimately we can say this process is statistically out-of control process.

Illustration 03: The net weight (in grams) of a yam powder product is to be monitored by $\bar{X}$ control chart and R control chart and $\bar{X}$ control chart and S control chart using a sample size of n=5. Data for 24 preliminary samples and particular quality control charts are shown in following outputs (Note: Exclude Sample No: 22).

Table 3: Net weight (in grams) of a yam powder product – Illustration 03

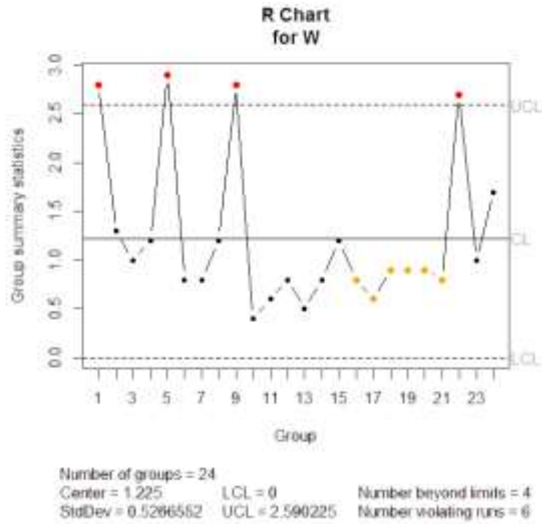
X1	X2	X3	X4	X5	$\bar{X}_i$	R_i	S_i
51.8	50	50.1	52.8	50.5	51.04	**2.8**	**1.217785**
51.2	50.8	51.3	51.4	52.1	51.36	1.3	0.472229
50.1	50.5	50.5	50	51	50.42	1	0.396232
50.3	50.8	51.5	51.4	51.2	51.04	1.2	0.49295
50	50.2	51.3	52.9	50.9	51.06	**2.9**	**1.154556**
51.1	51.2	50.8	51.4	50.6	51.02	0.8	0.319374
50.8	50.9	50.7	51.2	51.5	51.02	0.8	0.327109
50.2	51.2	50.8	51.4	51.2	50.96	1.2	0.477493
50.9	50	51	52.8	51.8	51.3	**2.8**	**1.053565**
50.2	50.3	50.2	50	50.4	50.22	0.4	0.148324
50.7	50.8	50.9	50.4	51	50.76	0.6	0.230217
50.7	50.8	50.4	50.2	51	50.62	0.8	0.319374
50.2	50.3	50.5	50.4	50	50.28	0.5	0.192354
50.3	50.2	50.8	50.2	50	50.3	0.8	0.3
50.6	50.4	50.2	51.1	51.4	50.74	1.2	0.497996
50	50	50.4	50.8	50.2	50.28	0.8	0.334664
51.1	51.2	50.9	50.6	50.7	50.9	0.6	0.254951
50.7	50.8	50.9	50	50.2	50.52	0.9	0.396232
50.8	50.2	50	50.3	50.9	50.44	0.9	0.391152
51	51.2	50.8	50.6	50.3	50.78	0.9	0.349285
50.2	50.9	50.5	51	50.6	50.64	0.8	0.320936
~~**50.1**~~	~~**50.2**~~	~~**50.8**~~	~~**50.7**~~	~~**70.0**~~	~~**54.36**~~	~~**19.9**~~	~~**8.74831**~~
52.7	50.2	50.3	50	50.9	50.82	**2.7**	**1.103177**
50.8	50.7	50.2	51.2	51.1	50.8	1	0.3937
50.3	50.2	50.1	51.2	51.8	50.72	1.7	0.746324

$$\bar{\bar{X}} = \frac{\sum_{i=1}^{m} \bar{x}_i}{m} \qquad \bar{R} = \frac{\sum_{i=1}^{m} R_i}{m} \qquad \bar{S} = \frac{\sum_{i=1}^{m} \bar{s}_i}{m}$$

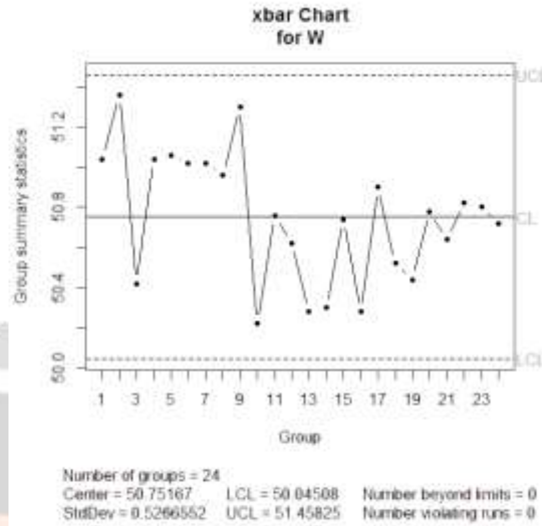
$$\underline{\bar{\bar{X}} = 50.7517} \qquad \underline{\bar{R} = 1.225} \qquad \underline{\bar{S} = 0.49542}$$

When the 24 sample ranges are plotted on the R control chart in Fig. 6 there is an indication of an out-of control condition. Since the R and S control chart indicates that process variability is out-of control, we may not go to discuss about the $\bar{X}$ control chart. Ultimately we can say this process is statistically out-of control process.

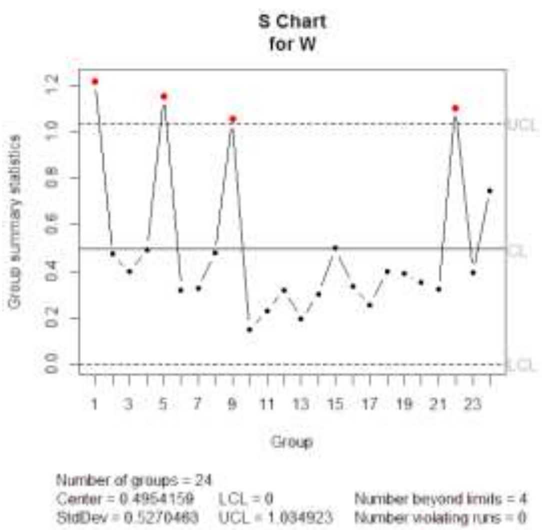
> q <- qcc(weight, type="R", nsigmas=3)



> q <- qcc(weight, type="xbar", nsigmas=3)



> q <- qcc(weight, type="S", nsigmas=3)



> q <- qcc(weight, type="xbar", nsigmas=3)

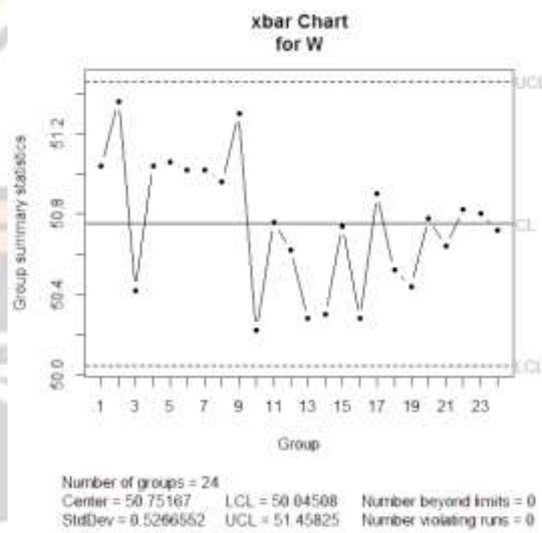


Figure 6: $\bar{X}$ *and* R/S *Control Chart* $(Illustration - 03)$

3. CONCLUSION

R and S Control Charts are quite a useful indication of how spread out the data is, but it has some serious limitations. This is because sometimes data can have outliers that are widely off the other data points. In these cases, the R and S Control Charts might not give a true indication of the spread of data. The R and S Control Charts can sometimes be misleading when there are extremely high or low values (outliers).

REFERENCES

[1]. Joiner, B.L. (1994): Fourth Generation Management. The New Business Consciousness. How the Evolution of Management and the Revolution in Quality are converging, and What it means for Business and the Nation. New York: McGraw-Hill.
[2]. Juran, J.M. (1999): How to think about Quality. In J.M. Juran, A.B. Godfrey, R.E. Hoogstoel, and E.G., Schilling (Eds.): Quality-Control Handbook. New York: McGraw-Hill.
[3]. Montgomery, D. C., 2004, Introduction to Statistical Quality Control, John Wiley, 759 pp.
[4]. Russell, R., Taylor, B., 2006, Operations Management - 5th Edition, Statistical Process Control, John Wiley & Sons, Inc. 10.
[5]. Sargut, K. U., Demirors, O., 2006, Utilization of statistical process control (SPC) in emergent software organizations: pitfalls and suggestions, Software Quality Journal 14(2), pp. 135-157. 11.
[6]. Zhelezarov, I., 2006, Statistical process control by the precontrol and digital measuring means with data output, Facta Univesitatis Series Mechanical Engineering, 4(1) pp. 55-62.