When Do We Use Subgrouped Data?

A Problem with Service Sector Average Charts

Donald J. Wheeler

When the data come along one value at a time we tend to put the data on a Chart for Individual Values (an *XmR* Chart). Since virtually all business and managerial data occur one value at a time, the primary chart for service sector data is the *XmR* Chart. Yet the original process behavior chart was the Average and Range Chart (also known as an X-bar and R Chart). While Average and Range Charts are somewhat more specialized than *XmR* Charts, they do have certain advantages which make them particularly well suited for specific applications. At the same time, their greater specialization imposes certain restrictions upon their use. The advantages, restrictions, and issues surrounding Average and Range Charts are the topics of this paper.

When it makes sense to arrange the data into subgroups you will gain sensitivity by doing so. To illustrate this aspect of an Average and Range Chart I will use the data from a bottle filling line. Each hour five bottles are taken from the line and the amount of product in each bottle is measured. These five fill weights are considered one subgroup and the average and range for each of these subgroups is plotted on the Average and Range Chart. Figure 1 shows the data for sixteen such subgroups.

Figure 1: Fill Weights for Sixteen Hourly Samples of Five Bottles

Hour	Weights					Average	Range
1	751	754	756	754	753	753.6	5
2	757	755	756	754	756	755.6	3
3	755	757	756	752	755	755.0	5
4	755	751	756	757	753	754.4	6
5	758	753	756	754	758	755.8	5
6	754	755	755	754	752	754.0	3
7	754	758	756	757	759	756.8	5
8	752	757	755	754	756	754.8	5
9	755	754	753	754	755	754.2	2
10	757	758	755	753	754	755.4	5
11	757	755	756	753	752	754.6	5
12	759	758	755	756	753	756.2	6
13	756	754	756	755	755	755.2	2
14	755	750	755	757	755	754.4	7
15	757	754	756	754	754	755.0	3
16	750	756	752	755	754	753.4	6

The Grand Average is 754.9, the Average Range is 4.5625. The scaling factors for creating an Average and Range Chart for subgroups of size five are $A_2 = 0.577$ and $D_4 = 2.114$. (Tables of these scaling factors are found in most books on SPC.) Thus, the limits for the Average Chart are:

$$\text{Grand Average} \pm A_2 \text{ Average Range} = 754.9 \pm 0.577\,(4.56) = 752.27 \text{ to } 757.53$$

On the other hand, had we placed the 80 individual fill weights on an *XmR* Chart we would have had an Average of 754.9 and an Average Moving Range of 1.987. This would have resulted in limits for the *X* Chart of:

Average ± 2.66 Average Moving Range = 754.9 ± 2.66 (1.99) = 760.2 to 749.6

Figure 2 compares the Average Chart with the Chart for Individual Values. Both charts tell the same story about this process. However, if the process changes, the Average Chart is likely to have a point outside the limits before the *XmR* Chart does.

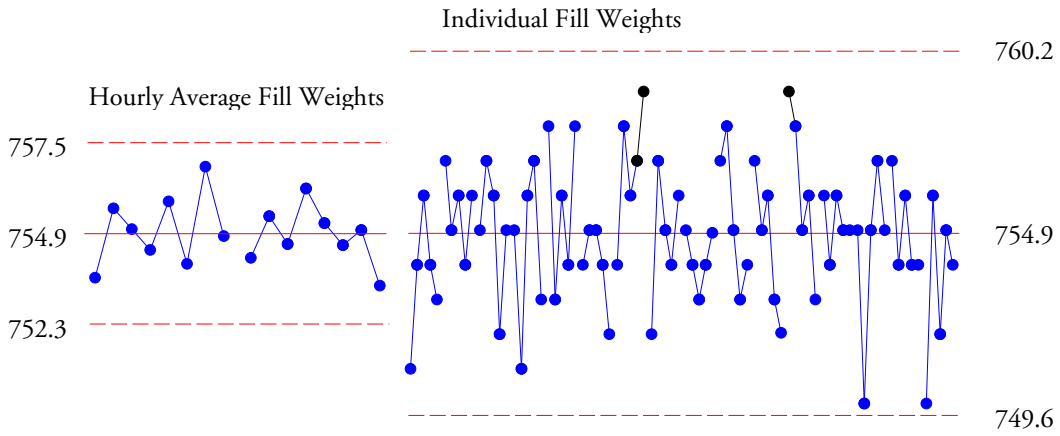

Figure 2: Average Chart and Individual Values Chart for the Fill Weight Data

This increased sensitivity of the Average Chart is dependent upon having a rational subgrouping. When you place two or more values together in a subgroup you are making a *judgment* that you consider these values to have been collected under essentially the same conditions. This means that the differences between the values within a subgroup should generally represent nothing more, and nothing less, than routine, background variation.

The first requirement for rational subgrouping is that of internal homogeneity for the subgroups.

One way to satisfy this requirement of internal homogeneity is to collect multiple measurements within a reasonably short time. While this is easy to do when making widgets, it becomes less feasible with continuous processes and administrative systems. In the example above we had one subgroup per hour and five pieces per subgroup. When we placed the five values from each hour together in one subgroup we were making a judgment that within each hour the five pieces could be considered to be essentially uniform.

This requirement of internal consistency means that you are willing to consider the variation within the subgroups as "background noise." The assumption that the values were collected under essentially the same conditions means that you are not concerned with the differences between the values. Thus these differences must be thought of as the background level of noise which is present in your system. And, as you may see from the computations, it is this background noise within the subgroups (as characterized by the average range) that is used to determine the limits for the Average and Range Chart.

The second requirement for rational subgrouping is that the variation within the subgroups must truly represent the reasonable and proper amount of background variation for the process.

In the previous example, when we computed the limits for the Average and Range Chart, we were using the average variation within the subgroups to set limits on how much variation we

should expect to see from subgroup to subgroup. We used the average variation among the five bottles collected each hour to determine how much variation should occur from hour to hour. And the only way to *judge* if this is appropriate is to know both the process and the way the measurements are made.

In other words, there has always been an element of judgment involved in subgrouping. Subgrouping cannot be automated. It is the conclusion of a *thought* process that combines the context for the data with the purpose of the process behavior chart and then organizes the data in such a way that the chart will answer the questions of interest.

So what about administrative and service sector data? There are many situations where people will tend to collect these data into subgroups. Two examples follow.

Figure 3 shows the daily sales figures for a department store over a three month period. These values are given to the nearest hundred dollars. Each week's worth of sales have been made into a subgroup and the weekly averages and ranges are included at the bottom of Figure 3

Figure 3: Daily Department Store Sales

Day	*Week 1*	*Week 2*	*Week 3*	*Week 4*	*Week 5*	*Week 6*	*Week 7*	*Week 8*	*Week 9*	*Week 10*	*Week 11*	*Week 12*	*Week 13*
Sun.	66	67	69	71	77	122	67	76	108	76	87	126	82
Mon.	44	46	56	86	67	76	65	85	77	60	64	55	68
Tues.	54	54	44	53	45	64	64	76	65	85	54	47	44
Wed.	108	107	100	92	112	100	89	87	98	71	107	103	121
Thurs.	78	109	88	70	90	70	89	103	100	100	99	113	100
Fri.	107	86	127	111	65	85	109	88	74	111	114	75	100
Sat.	150	142	149	160	189	139	163	151	158	171	159	180	180
Averages	86.7	87.3	90.4	91.9	92.1	93.7	92.3	95.1	97.1	96.3	97.7	99.9	99.3
Ranges	106	96	105	107	144	75	99	75	93	111	105	133	136

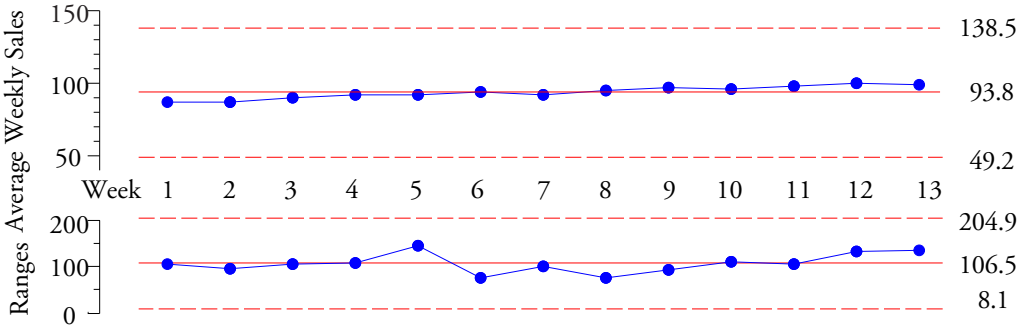

Figure 4: Average and Range Chart for Daily Department Store Sales

The Grand Average is 93.84, and the Average Range is 106.5. The scaling factors for subgroups of size $n = 7$ are: $A_2 = 0.419$, $D_3 = 0.076$, and $D_4 = 1.924$. The limits for the Average Chart are:

$$\bar{\bar{X}} \pm A_2 \bar{R} \;=\; 93.84 \;\pm\; 0.419\,(106.5) \;=\; 49.21 \;\text{ to }\; 138.46$$

while the Range Chart limits are:

$$D_3 \bar{R} \;=\; 0.076\,(106.5) \;=\; 8.1 \quad \text{and} \quad D_4 \bar{R} \;=\; 1.924\,(106.5) \;=\; 204.9$$

Does the subgrouping of the Daily Department Store Sales shown in Figure 3 satisfy the two requirements of rational subgrouping? Are the subgroups internally homogeneous? Does the

variation within the subgroups provide the proper yardstick for measuring the variation from subgroup to subgroup?

Even a cursory reading of the data in Figure 3 shows that there are substantial differences in the sales volumes of different days of the week, with Mondays and Tuesdays being low and Saturdays being highest. Therefore, when the seven days of one week are subgrouped together we have a collection of unlike things. This represents a violation of the first requirement of rational subgrouping.

In addition, the variation within the subgroups is the day-to-day variation, while the variation between the subgroups is the week-to-week variation. Since the day-to-day variation is a completely different type of variation than the week-to-week differences, it is unreasonable to expect that the variation within the subgroups will provide an appropriate yardstick for setting the limits on the Average Chart.

Therefore, the weekly subgroups shown in Figure 3 violate both requirements of rational subgrouping. While it is possible to perform the computations and to create a chart using the "subgrouping" of Figure 3, the failure to organize the data into *rational* subgroups will result in a chart that is useless.

How then can we track these sales over time? Could we plot these 91 values in a running record of Daily Sales? Yes. Should we turn this running record of 91 Daily Sales values into an *XmR* Chart? Probably not.

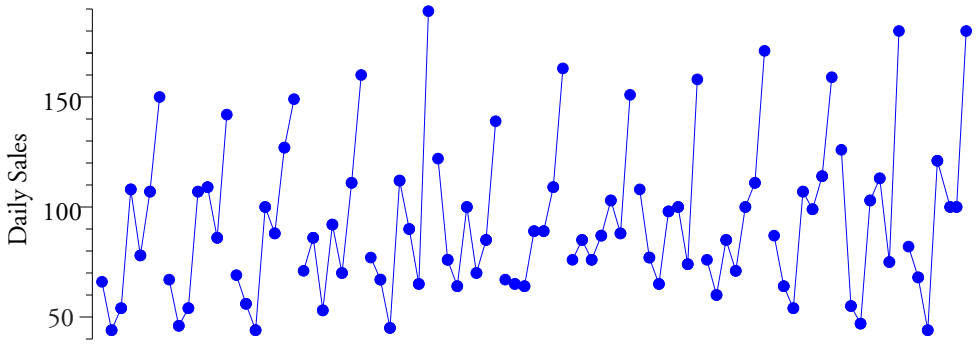

Figure 5 Running Record for Daily Department Store Sales

If we placed the 91 values from Figure 3 on an *XmR* Chart, the strong daily cycle of sales would inflate the moving ranges, which would in turn inflate the limits. The principles of rational subgrouping require successive values on an *XmR* Chart to be collected under conditions that are, at least most of the time, reasonably similar.

Before we can place data on an XmR Chart successive values must be logically comparable.

When your data display a recurring pattern this requirement is not satisfied. The name given to recurring patterns like the one in Figure 5 is seasonality. (To discover how to deal with seasonality when placing data on a process behavior chart see my book *Making Sense of Data*.)

Could we plot the weekly average daily sales on an *XmR* Chart? Yes. This will probably be the most satisfying way of tracking these data over time. Using the data in Figure 3, the overall Average is still 93.84, and the Average Moving Range of the weekly averages is 1.52, resulting in the limits shown in Figure 6. While the running record on the *X* Chart in Figure 6 is the same as that of the Average Chart in Figure 4, the story told by these two charts is considerably different. The Average Chart is less sensitive because the subgrouping used there was not rational.

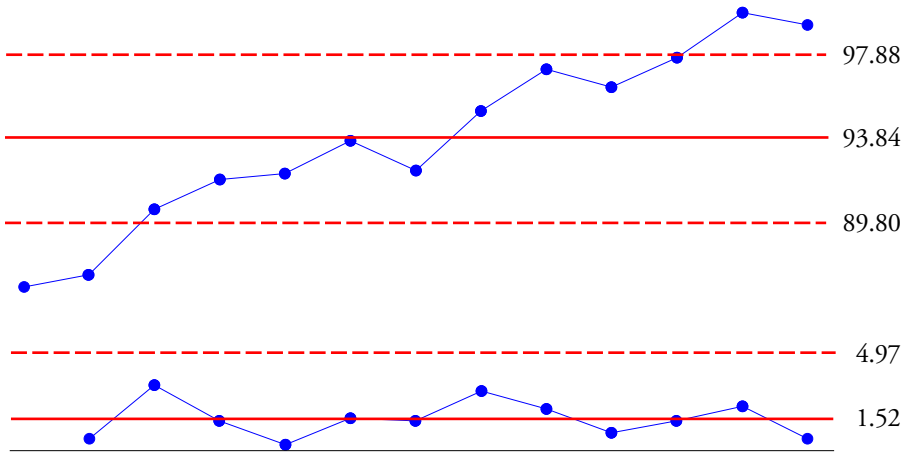

Figure 6 *XmR* Chart for Weekly Average Daily Sales

Figure 7 provides a second example of how administrative data are often arranged into subgroups. There you will find the quarterly sales for each of six regions for the past five years. These values have been arranged into 20 subgroups of size 6 so that each quarter is a subgroup.

Figure 7: Quarterly Sales for Five Years

Quarter	Region A	Region B	Region C	Region D	Region E	Region F	Averages	Ranges
1	924	1,412	1,056	539	397	431	793.2	1015
2	928	1,280	1,048	558	391	470	779.2	889
3	956	1,129	1,129	591	414	439	776.3	715
4	1,222	1,181	1,073	556	407	431	811.7	815
5	748	1,149	1,157	540	415	471	746.7	742
6	962	1,248	1,146	590	442	496	814.0	806
7	983	1,103	1,064	606	384	506	774.3	719
8	1,024	1,021	1,213	643	448	573	820.3	765
9	991	1,085	1,088	657	441	403	777.5	685
10	978	1,125	1,322	602	366	440	805.5	956
11	1,040	910	1,256	596	470	371	773.8	885
12	1,295	999	1,132	640	426	405	816.2	890
13	765	883	1,352	691	445	466	767.0	907
14	1,008	851	1,353	723	455	536	821.0	898
15	1,038	997	1,466	701	363	551	852.7	1103
16	952	878	1,196	802	462	670	826.7	734
17	1,041	939	1,330	749	420	588	844.5	910
18	1,020	834	1,003	762	454	699	795.3	566
19	976	688	1,197	807	447	743	809.7	750
20	1,148	806	1,337	781	359	702	855.5	978

The Grand Average for the subgroups of Figure 7 is 803. The Average Range is 836. The Average and Range Chart for these 20 subgroups is shown in Figure 8. Before you attempt to interpret Figure 8 you need to answer two questions regarding the subgrouping used. These questions are: What type of variation occurs *within* the subgroups in Figure 7? What type of variation occurs *between* subgroups in Table 7?

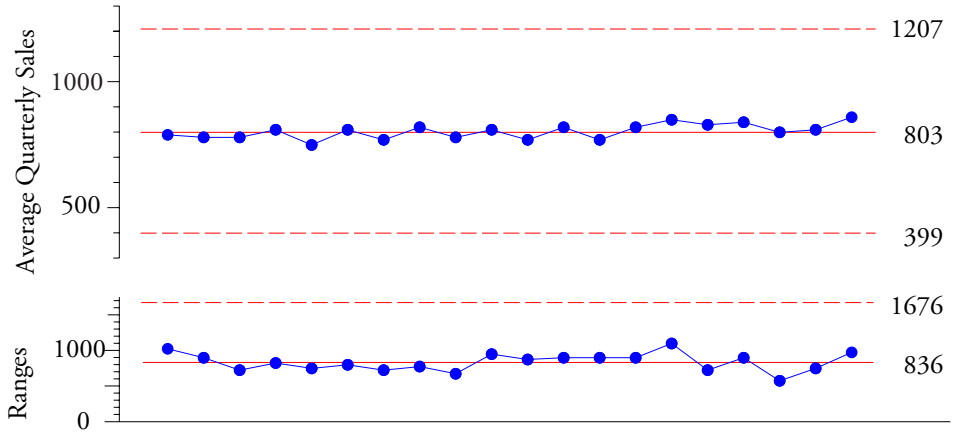

Figure 8: Average and Range Chart for Quarterly Sales

What is the problem with the subgrouping in Figure 7? Are the subgroups internally homogeneous? It is the region-to-region variation that occurs within the subgroups in Figure 7, while the quarter-to-quarter variation shows up between the subgroups. Is the region-to-region variation the proper yardstick for characterizing the variation from quarter to quarter? Would it not make more sense to acknowledge that the six regions are different, and then to track each region separately on an *XmR* Chart?

THE STRATIFICATION OF SUBGROUPS

Subgroups are said to stratified when each subgroup contains systematic differences in addition to routine variation. One of the ways this happens is when we subgroup unlike things together. The problem of stratification will be apparent on the Average and Range Chart because it will result in limits that look too wide for the graph. This phenomenon may be seen in both Figures 4 and 8. When the computations for an Average and Range Chart result in limits that look far too wide for the graph, check to see if unlike things have been grouped together. In Figure 4 it is the unlike daily sales that are subgrouped together. In Figure 8 it is the regions that are different that have been subgrouped together. When looking at these charts we see that both the Averages and the Ranges are hugging the central line. Whenever 15 or more successive values on the Range Chart fall within one sigma of the central line, you should check for possible stratification of the subgroups.

The Average and Range Chart, for all of its power and sensitivity, requires data which contain multiple measurements collected under essentially the same conditions. This requirement is rarely satisfied in service and administrative applications where data tend to come along one value at a time. This means that it will be extremely difficult to find ways of using Average and Range Charts to directly monitor service sector or administrative data. As outlined above, there are many ways to arrange data into what look like subgroups, but not all of these will result in charts that are appropriate, useful, or meaningful. If the subgrouping is not rational, then the chart will not work correctly, and you will be in danger of misleading yourself and others.

Computations simply cannot correct the defects of bad subgrouping. Judgment is always required in order to organize the data into rational subgroups.