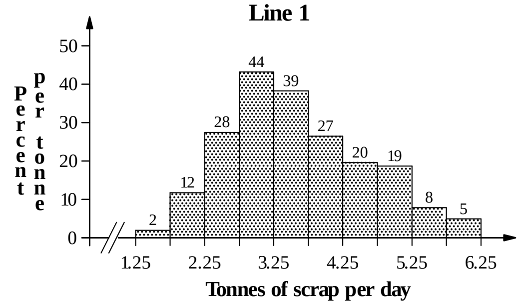
Assignment 4 Outline Solution

A4 – 1. (a) From the data given in the statement of the question, to prepare for the construction of the required histograms we tabulate the relative frequencies and bar heights as shown at the right.

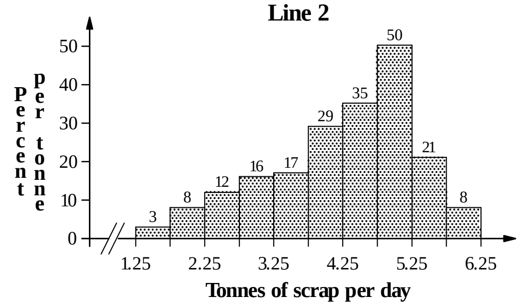
The histograms (shown below the table) for these data both cover the same interval of values (1.25 to 6.25 tonnes of scrap per day) and both are roughly 'mound-shaped'; however, the histogram for Line 1 is skewed to the *right*, whereas the one for Line 2 is skewed to the *left*..

Class Interval	Line 1 Frequency No.	Line 1 Frequency %	Line 1 Bar Height	Line 2 Frequency No.	Line 2 Frequency %	Line 2 Bar Height
[1.25, 1.75)	2	1.0	1.96	3	1.5	3.02
[1.75, 2.25)	12	5.9	11.76	8	4.0	8.04
[2.25, 2.75)	28	13.7	27.45	12	6.0	12.06
[2.75, 3.25)	44	21.6	43.14	16	8.0	16.08
[3.25, 3.75)	39	19.1	38.24	17	8.5	17.09
[3.75, 4.25)	27	13.2	26.47	29	14.6	29.15
[4.25, 4.75)	20	9.8	19.61	35	17.6	35.18
[4.75, 5.25)	19	9.3	18.63	50	25.1	50.25
[5.25, 5.75)	8	3.9	7.84	21	10.6	21.11
[5.75, 6.25)	5	2.5	4.90	8	4.0	8.04
	204	100.0		199	99.9	

Line 1



Line 2



(b) In the table above at the rght, we take the sum of the products of each interval mid-point times the corresponding frequency, all divided by n, the sum of the frequencies for the relevant Line; this gives:

Line 1: $732.5/204 \simeq 3.59 \simeq 3.6$ tonnes per day;

Line 2: $845.0/199 \simeq 4.25$ tonnes per day.

(c) From (b), about 732 tonnes in a year for Line 1;
about 845 tonnes in a year for Line 2.

i.., Line 2 produces more scrap by about 113 tonnes.

(d) Reasoning as in (b), we find averages of:

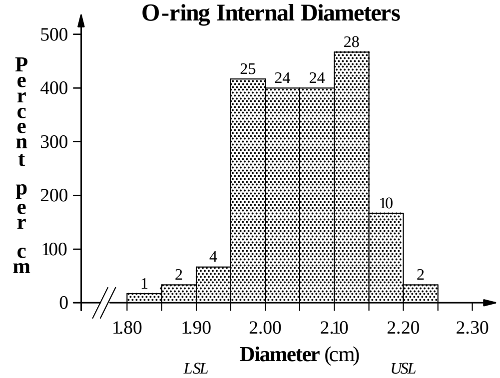
Line 1, Part *A*: about $415.5/135 \simeq 3.08$ tonnes per day;
Part *B*: about $317.0/69 \simeq 4.59$ tonnes per day;

Line 2, Part *A*: about $162.0/55 \simeq 2.95$ tonnes per day;
Part *B*: about $683.0/144 \simeq 4.74$ tonnes per day.

(e) From these four scrap tonnage rates, we see that the two lines have quite *similar* rates for the two parts *A* and *B* considered *separately* (Line 2 is a little better than Line 1 for *A*, a little worse for *B*); thus, the unfavourable view of Line 2 from (c) essentially *disappears* under this stratified analysis of the scrap rates. A more precise comparison requires additional information – the *number* of each part produced by each line; the calculation above of the four approximate averages assumed that these numbers are proportional to the amount of scrap, a reasonable assumption in the circumstances but probably not entirely correct.

A4 – 2. (a) The histogram is shown at the right; the interval width of 0.05 cm is more appropriate in this situation than the other reasonable choice of 0.06 cm.

O-ring Internal Diameters



(b) From the 120 diameters ordered from smallest to largest, we find:

average = 2.0505 cm
median = 2.05 cm
mode = 2.00 cm.

(c) We also find:

range = 2.22 – 1.83 = 0.39 cm
s.d. = 0.0764 cm
IQR = 2.11 – 1.99 = 0.12 cm.

(d) The boxplots are shown and discussed overleaf on page 0.12b.

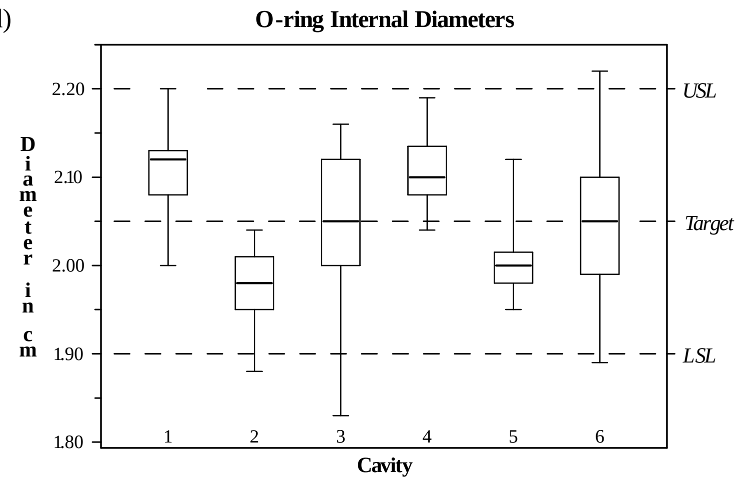
(e) The sampling should be done *by cavity* with *equal* sample sizes per cavity.

(f) The data recording sheets should keep the data *by cavity* and also retain the observations in the *order* in which the O-rings are produced; this would assist with identifying and reducing source(s) of variation.

(*continued overleaf*)

Assignment 4 Outline Solution (continued 1)

A4 – 2. (cont.) (d)



These boxplots show several indications of poor process performance:

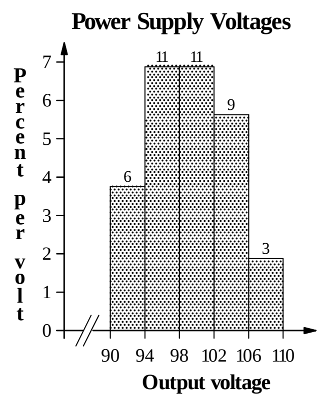
– cavities 1, 2, 4 and 5 need to be adjusted to get them on target;

– cavities 3 and 6 are reasonably well on target but show too much variation.

The aim would be to have *each* cavity on target and with a variation like that of cavity 4 or, preferably, cavity 5.

A4 – 3. From the data given in the question, we tabulate the relative frequencies and bar heights as given below; we then construct the required histogram shown at the right.

Class Interval	Frequency No.	%	Bar Height
[90, 94)	5	15	3.750
[94, 98)	11	27½	6.875
[98, 102)	11	27½	6.875
[102, 106)	9	22½	5.625
[106, 110)	3	7½	1.875
	40	100	



We see from the histogram that the distribution of output voltages is somewhat *uniform* with rather abrupt edges; a more *normal* distribution might be expected. If voltage is a critical characteristic in the operation of the power supplies, this shape for the distribution may indicate that there is a problem with the manufacturing process of the power supplies.

A4 – 4. A process is *in control* (or it is a *stable process*) when its pattern of variation, assessed by measurement of suitable process characteristic(s), is the *same over time*.

- We can assess process stability by examining histograms of the measurements of the process output(s) for successive time periods; each period should be long enough to provide sufficient data to give a reliable histogram.
  - For a stable process, the shape, centre, and spread of the histograms do not change appreciably from one time period to the next.
  - A stable process (*i.e.*, a process in control) is *predictable* in that the aggregate characteristic(s) of its output remain the *same* over time.
- **NO**: a process that is using only a relatively small fraction of its tolerance could be out of control but still producing 100% conforming product – recall the *Japanese Control Chart* video described in Figure 11.25 of the Course Materials. Further, a process that is *in control* but well off target might be producing *no* conforming product.

A4 – 5. When a control chart is to be used for process *adjustment*, the centre line and control limits are based on specifica-'tions (or standards).

Hence: centre line is 25°C;

control limits are $\pm\, 3\dfrac{\sigma}{\sqrt{n}} = \dfrac{3 \times 0.40}{\sqrt{5}} = 0.537°\text{C}$ so $\begin{cases} UCL = 25.54°\text{C}; \\ LCL = 24.46°\text{C}. \end{cases}$

A4 – 6. When a control chart is to be used for *process improvement*, observations which are clearly atypical should be removed *before* the data are used to calcuate control limits. For the observations given [which are *averages* of *3* (hopefully independent) breaking strength measurements], the *13*th value (*viz*. 7.59 p.s.i.) is sufficiently different (on the *low* side) from the others to be removed.

Hence, based on the other 14 averages: $\bar{\bar{y}} = \dfrac{\Sigma \bar{y}_i}{14} = \dfrac{169.87}{14} = 12.134$ p.s.i.;

also, based on the historical standard deviation of 1.2 p.s.i. and the sample size of 3:

(*continued*)

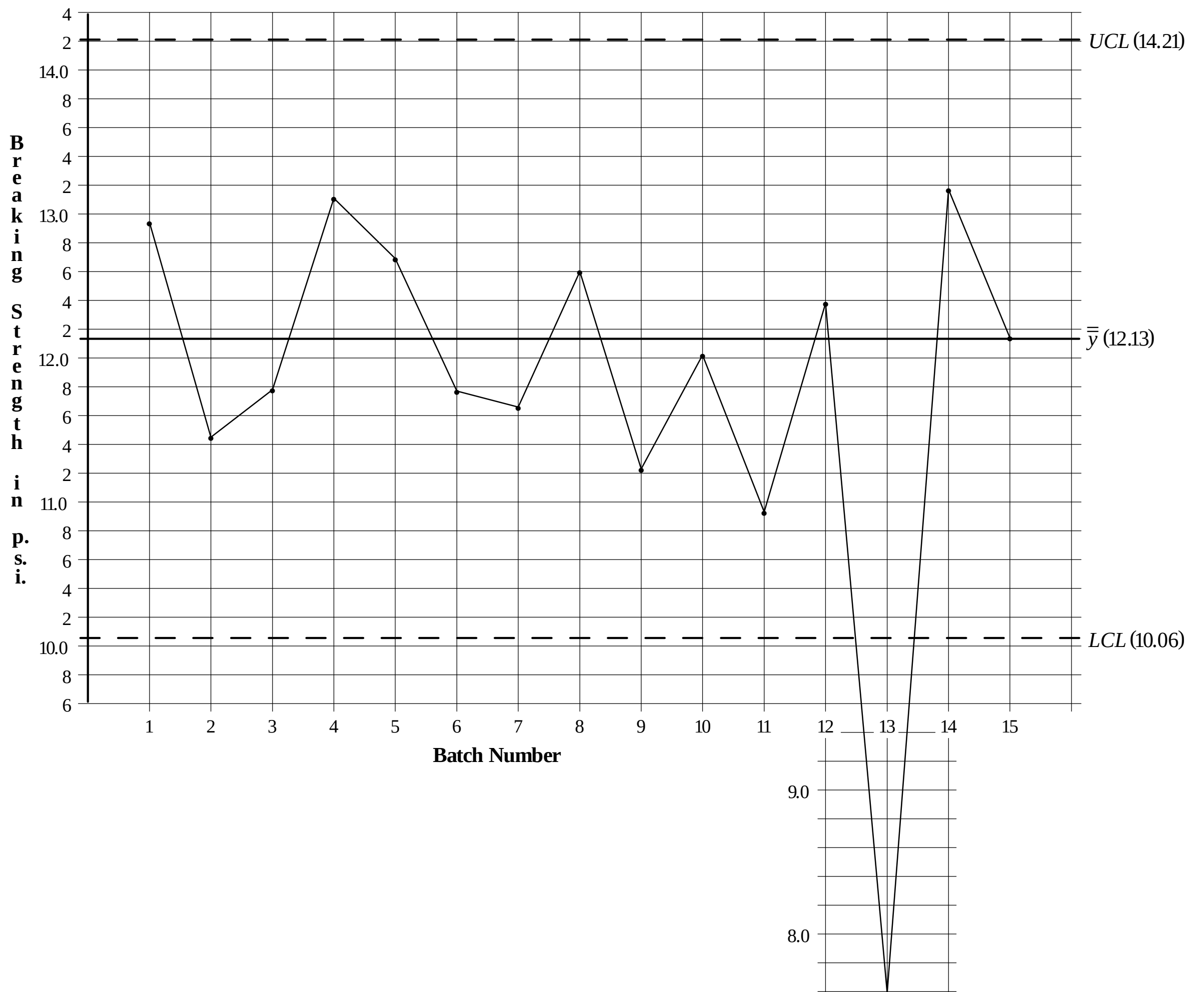
Assignment 4 Outline Solution (continued 2)

A4 - 6. (a) **(cont.)** control limits are $\pm 3\frac{\sigma}{\sqrt{n}} = \frac{3 \times 1.2}{\sqrt{3}} = 2.08$ p.s.i. so $\begin{cases} UCL = 14.21 \text{ p.s.i.;} \\ LCL = 10.06 \text{ p.s.i.} \end{cases}$

We see from the control chart below that batch #13 *was* out of control, and so we *are* correct in *ex*cluding it from the calculation of $\overline{\overline{y}}$ and the control limits.

NOTE: The standard deviation of the 14 averages (*ex*cluding #13) is 0.7067, which estimates $\sigma/\sqrt{3}$; thus, the 14 averages give an estimate of σ of $\sqrt{3} \times 0.7067 = 1.224$ p.s.i., in good agreement with the historical value of 1.2 p.s.i.

Control Chart for Average Breaking Strength of Ceramic Insulators (n = 3)



(b) The *one point out* signal identifies batch #13 as indicating that the process is out of control, as noted above in (a); there is *no* signal for a *run of nine*.

The *action* recommended is to work with the process at the time of batch #13 to try to identify one or more *special causes*. The longer the time that is allowed to elapse since the out-of-control signal occurred, the *harder* it usually is to identify its cause(s).

(c) We assume that insulator breaking strength in this situation, represented by the random variable B, can be modelled by a normal distribution; we estimate the parameters of this model by the relevant process characteristics: *i.e.*, estimate μ by $\overline{\overline{y}} = 12.134$ p.s.i. and σ by the historical s.d. of 1.2 p.s.i., so that: $B \dot{\sim} N(12.134, 1.2)$.

Then: $\Pr(B < 10) = \Pr[N(0,1) < \frac{10 - 12.134}{1.2}] = \Pr[N(0,1) < -1.7780]$

$= 1 - 0.962\,296 = 0.037\,704 \simeq 3.8\%$;

i.e., a little under 4% of the insulators do not meet the breaking strength specification.

(*continued overleaf*)

Assignment 4 Outline Solution (continued 3)

A4 – 7. (a) We calculate the following numerical summaries of the three data sets:

DATA SET	$\Sigma\overline{y}_i$	$\Sigma\overline{y}_i^{\,2}$	$\overline{\overline{y}}$	$s_{\overline{y}}$	$3s_{\overline{y}}$	*LCL*	*UCL*
A	231.651	2,683.4575	11.5826	0.135 369	0.406 107	11.177	11.989
B	230.328	2,652.7206	11.5164	0.094 943	0.284 830	11.232	11.801
C	232.072	2,693.3874	11.6036	0.164 905	0.494 715	11.109	12.098.

We see (in the *fifth* column of the table) that the standard deviations of the three (data sets of) averages *do* have the values given in the statement of the question.

NOTE: Because the 15 values in each data set are all *averages* of $n = 4$ observations, their standard deviation ($s_{\overline{y}}$) estimates *directly* $\sigma/\sqrt{n}$; hence, the control limits in the table above are calculated as $\overline{\overline{y}} \pm 3s_{\overline{y}}$.

(b) The three control charts for averages are shown on the facing page 0.12e.

(c) Looking at the charts, we see the following:

- no *individual* point is out of control in any of the three data sets (but see the Note below);
- in **A**, starting with the *11*th observation, there is a run of 10 points *above* the centre line, indicating loss of control by the *19*th observation, the ninth consecutive point above the centre line.
- in **C**, the chart starts with a run of 11 points *below* the centre line, indicating loss of control by the *9*th observation;
  – it is likely that the *next* (*i.e.*, the *21*st) observation in **C** will be the *seventh* point in a run currently comprising six consecutive points *above* the centre line, suggesting that the process may again be going out of control but in the *opposite* direction from the initial situation.

In summary: **B** is in control;
A shows a sudden upward shift (at observation 11 or 12);
C shows a gradual drift upwards.

NOTE: Because neither **A** nor **C** appears to be in control, they are *not* suitable data sets from which to calculate the centre line and control limits for a control chart.

The problems with these two data sets are apparent from changes in their averages and/or standard deviations when the observations are considered *separately* for each half (*i.e.*, the *first* 10 *vs* the *last* 10 observations) of each data set, as follows:

DATA SET	FIRST HALF $\overline{\overline{y}}$	FIRST HALF $s_{\overline{y}}$	SECOND HALF $\overline{\overline{y}}$	SECOND HALF $s_{\overline{y}}$	BOTH HALVES $\overline{\overline{y}}$	BOTH HALVES $s_{\overline{y}}$
A	11.4811	0.091 220	11.6840	0.086 583	11.58	0.135
C	11.4910	0.069 674	11.7162	0.156 152	11.60	0.165 .

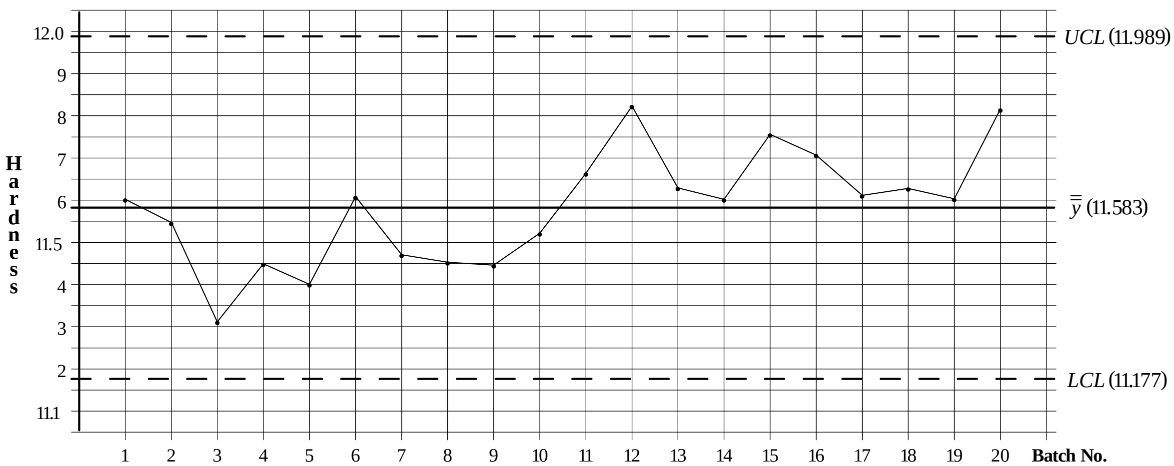
Thus, for **A**, each half of the data set is *less* variable than the overall data set; the higher overall standard deviation is a consequence of the increase in average for the second half of the data. For **C**, the second half of the data set has a higher average *and* greater variation than the first half.

(*continued*)

Assignmentt 4 Outline Solution (continued 4)

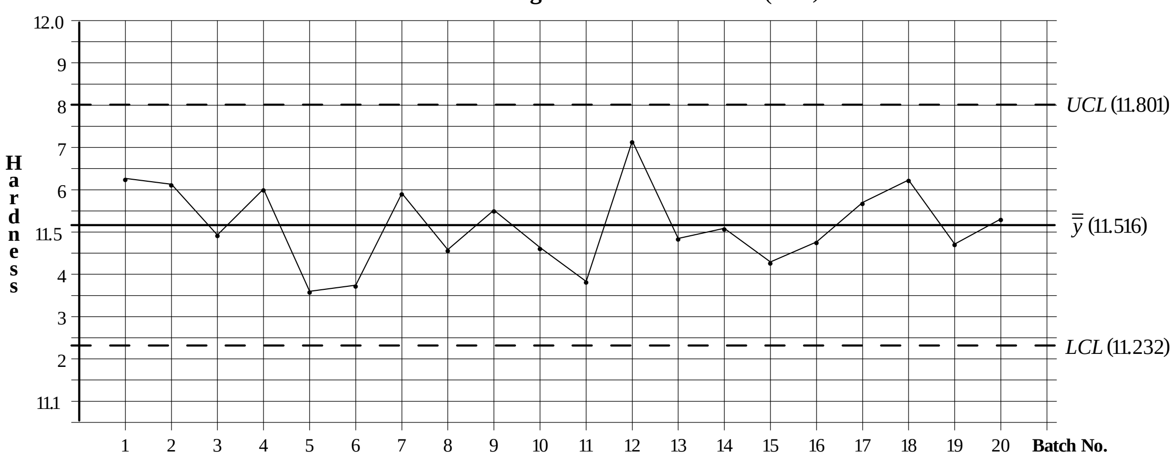
A4 – 7. (b)
(cont.)

Control Chart for Average Hardness of Tablets (n = 4)



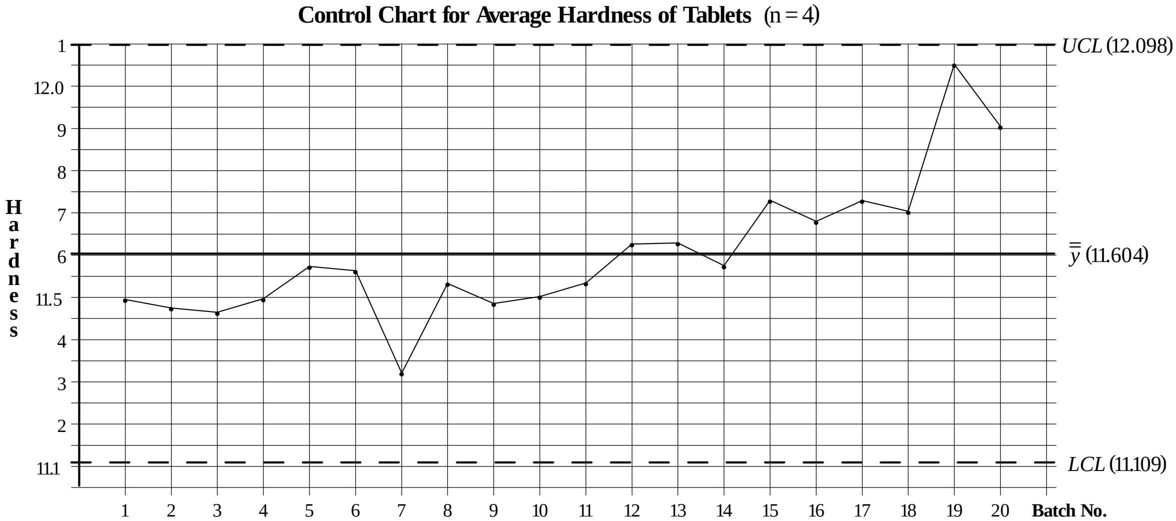
Data Set A

Control Chart for Average Hardness of Tablets (n = 4)



Data Set B

Control Chart for Average Hardness of Tablets (n = 4)



Data Set C