SUPPLEMENTARY EXERCISES, CHAPTER 1

INTRODUCTION TO DATA ANALYSIS

S1.1. Does it make a difference on what day of the week your car was made? Quality personnel for one of the major car manufacturers produced a quality index for each of 46 new cars of the same make and model manufactured over a two week period. The index took into account the number and severity of defects, and the quality of the fit and finish. The larger the index, the higher the quality. Figure 1 shows a plot of the quality index values stratified by the day of the week the car was produced.

a. What should the investigators have checked about the process prior to using this plot?

b. Assume the check you described in part a has been done and that the result was satisfactory. Analyze the data using the stratified plot. In particular, address the components of variation in the quality index and tell whether it appears there is a difference in the quality of cars produced on different days.

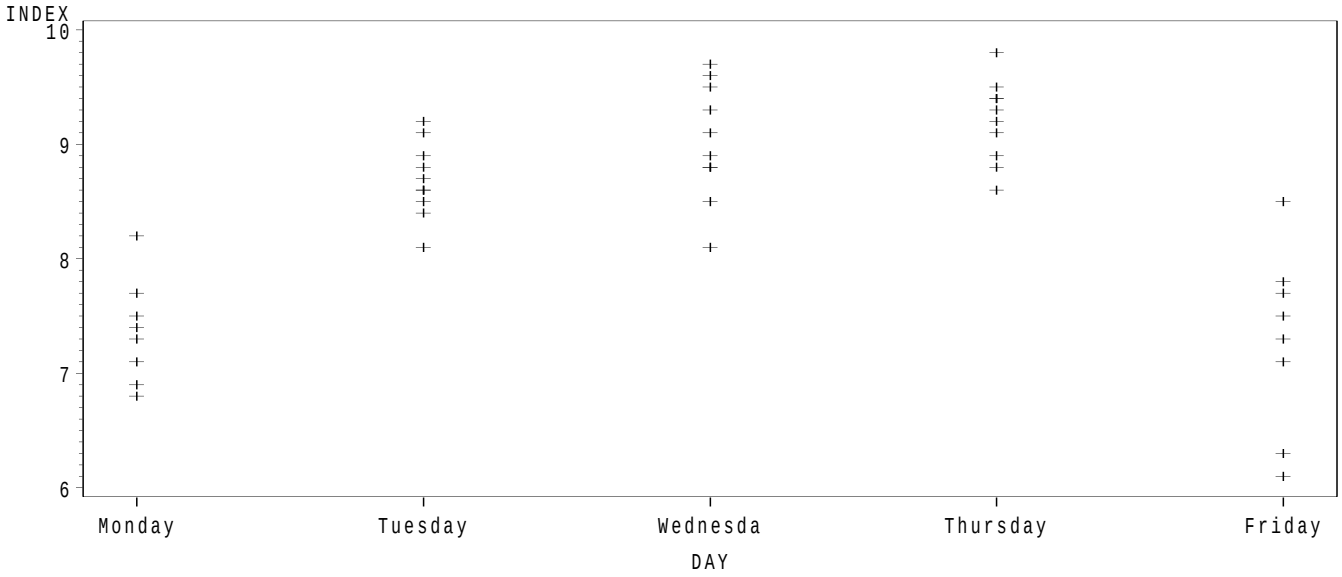

Figure 1: *Stratified plot of quality index by day, car quality data, Problem S1.1.*

S1.2. Zeolites are aluminum and silicon crystals which are used as sieves and catalysts in many industries-especially petroleum refining. Figure 2 shows frequency histograms of relative yields, in percent, of 30 samples taken from each of five consecutive production batches of zeolites.

a. Is the process stationary? Why or why not?

b. Suppose all five batches' worth of zeolite yields were plotted on a single histogram. Which of the histograms in Figure 3 would result?

S1.3. Three regulars at a fitness center weigh themselves once per week. The deviations from their ideal target weights for the last six weeks are:

	Week					
Weighee	1	2	3	4	5	6
Al	−2.9	1.7	−2.2	0.1	1.2	−2.7
W	−1.0	0.7	−1.8	−0.9	1.7	−2.1
Ralph	−0.5	−0.5	0.2	−0.1	0.1	0.2

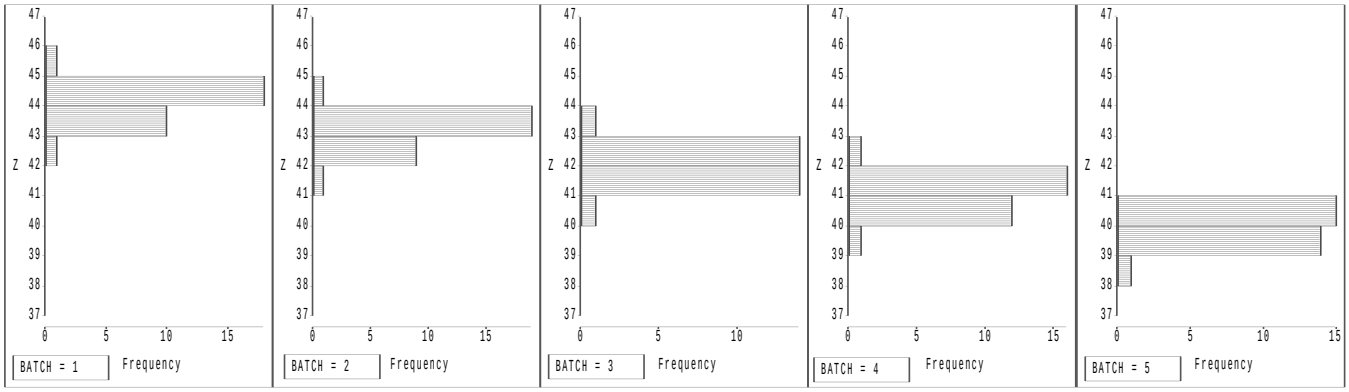

Figure 2: *Frequency histograms of relative yields, in percent, of five consecutive production batches of zeolites.*

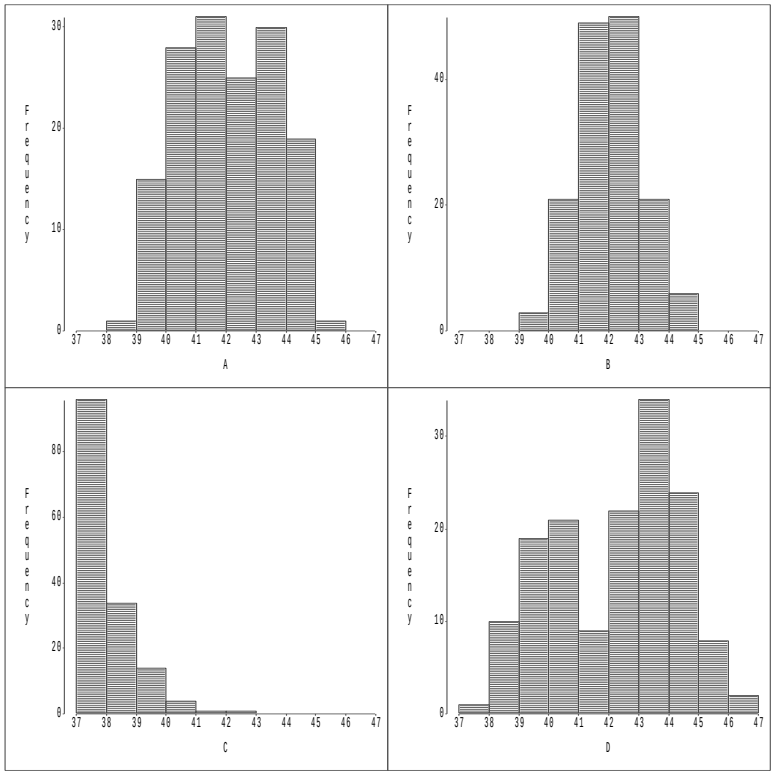

Figure 3: *Four possible Frequency histograms of relative yields, in percent, of production batches of zeolites.*

Assuming the data are from stationary processes, construct a plot to assess the pattern of between and within variation. Analyze the plot.

S1.4. Table 1 lists the postwar olympic winning times in the 100 meter dash.

a. Draw an appropriate graphical summary of these data. Explain why your summary is appropriate. What conclusion do you draw from your summary?

b. Jesse claims that the data are well summarized by their mean $\overline{y} = 10.1057$ and standard deviation $s = 0.2205$. Do you agree? Give your reasons.

S1.5. A fastener manufacturer wants to conduct a gage R&R study to assess the performance of a gage for measuring the lengths of 3/8" plastic rivets. To do so, they have three operators measure the same rivet ten times in succession using the same gage. The data are shown in Table 2.

Year	Time	Year	Time
1948	10.3	1976	10.16
1952	10.4	1980	10.25
1956	10.6	1984	9.99
1960	10.2	1988	9.92
1964	10.0	1992	9.96
1968	9.95	1996	9.84
1972	10.14	2000	9.87

Table 1: *Postwar olympic winning times in the 100 meter dash.*

	Operator		
Order	1	2	3
1	.3730	.3741	.3769
2	.3739	.3752	.3769
3	.3744	.3748	.3766
4	.3747	.3751	.3762
5	.3750	.3744	.3759
6	.3745	.3749	.3756
7	.3755	.3753	.3758
8	.3756	.3749	.3752
9	.3759	.3752	.3749
10	.3760	.3747	.3747

Table 2: *Data for exercise S1.5.*

What is the first graph you will produce to examine these data? Produce the graph and draw appropriate conclusions, giving your reasons.

S1.6. A gage R & R study was conducted on a weighing process. Three operators weighed a 1 pound weight five times each using the same scale. The data are (deviations, in pounds, from 1 pound):

Operator		
1	2	3
−0.007	−0.004	0.014
−0.008	0.005	0.022
−0.001	0.003	0.007
0.002	0.006	0.020
0.005	−0.001	−0.002

Assuming the data are from a stationary process, analyze the variation in the measurement process in terms of repeatability and reproducibility. Be sure to include an appropriate graph.

S1.7. Swimco Company's quarterly sales of swimming pools and accessories over a two year period are:

Year	Quarter	Sales ($1000s)
1	1	12
	2	147
	3	103
	4	28
2	1	14
	2	166
	3	116
	4	33

(a) Construct the most appropriate initial plot of these data from among those discussed in class.

(b) Choose the most appropriate moving average to remove the seasonal variation from these data.

(c) Apply the moving average chosen in (b) to these data, and use a plot to demonstrate that the seasonality has been removed.

S1.8. A metal stamping process consists of four presses simultaneously making metal clips for the automotive industry. The inclination angle of the clips (in degrees) determines the pressure they exert when placed in an assembly. The target angle is 10 degrees. Figure 4 displays the inclination angles of a set of clips sampled every 15 minutes from production plotted versus press number.

(a) Describe the pattern of variation you see in the plot.

(b) What should have been checked before using the plot in Figure 4? How would it be checked?

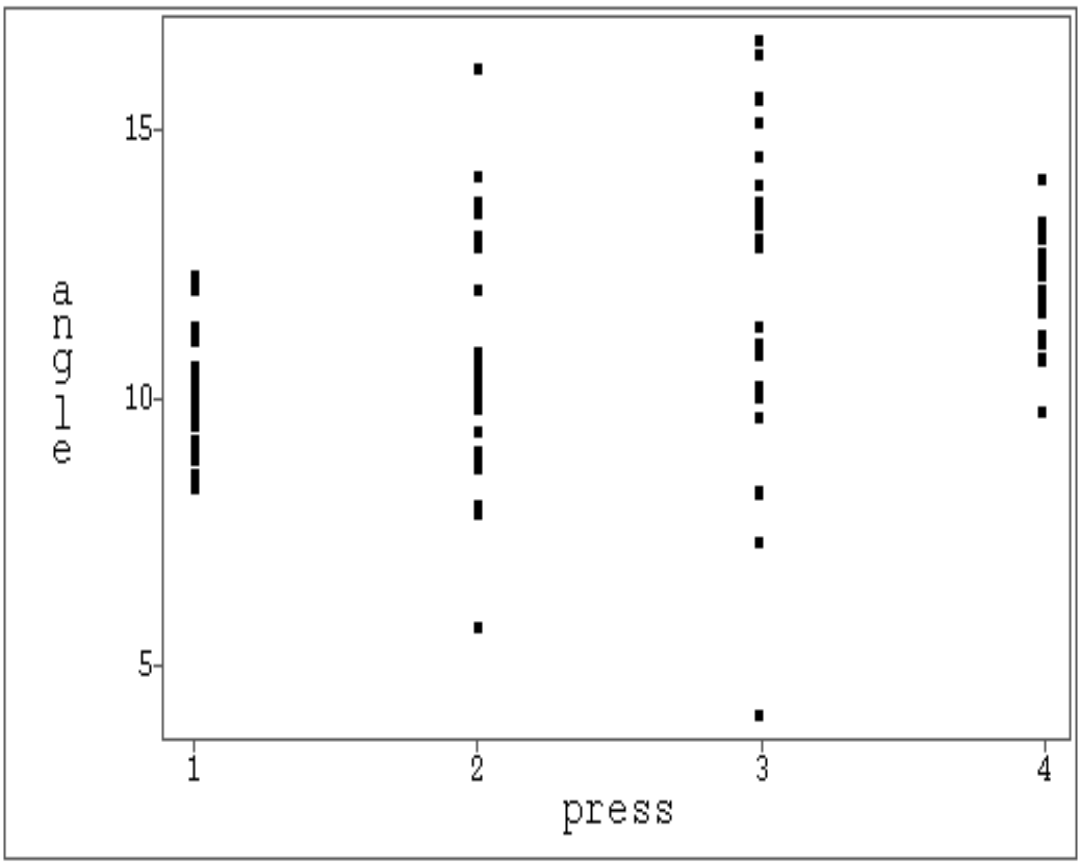

Figure 4: *Plot of Angle versus Press, problem 4.*

S1.9. Two years' worth of quarterly sales (in $ million) for Samco Corp. are shown in the table.

Quarter	Sales	Quarter	Sales
1	18.1	5	19.2
2	20.3	6	19.7
3	15.5	7	16.8
4	12.6	8	13.3

(a) Draw an appropriate plot of the data.

(b) Pick an appropriate moving average to eliminate the seasonality in the data. Justify your choice.

(c) Compute the moving average for the last four quarters, and plot it on the graph you drew in part a. Does it eliminate the seasonality?