SUPPLEMENTARY EXERCISE SOLUTIONS, 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?

ANS: *Stationarity.* **(5 points)**

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.

ANS: *There is substantial between-day variation, mainly attributable to lower quality scores for cars produced on Mondays and Fridays.* **(5 points)** *There is also noticeably greater variation (i.e., less consistency) in the quality index among cars produced on Fridays than on any other day of the week. The within variation for the other days is roughly equal.* **(5 points)** *The message here is that there will be more quality problems with cars produced on Mondays and Fridays.* **(5 points)**
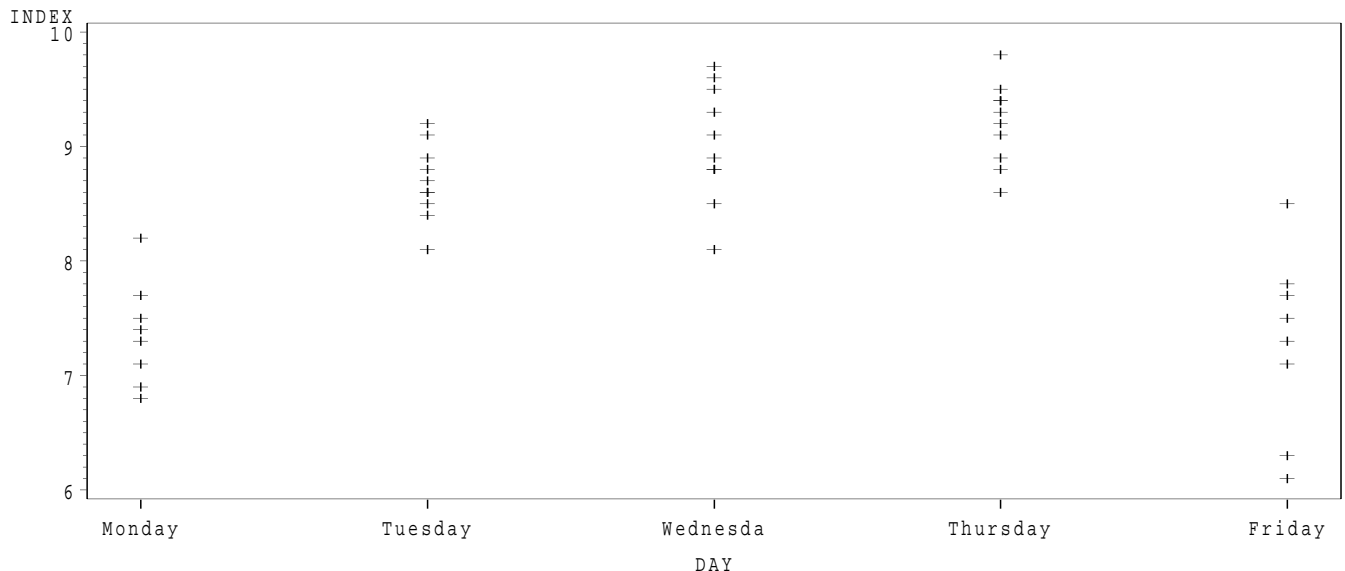

Figure 1: *Stratified plot of quality index by day, car quality data, Problem S1.1.*

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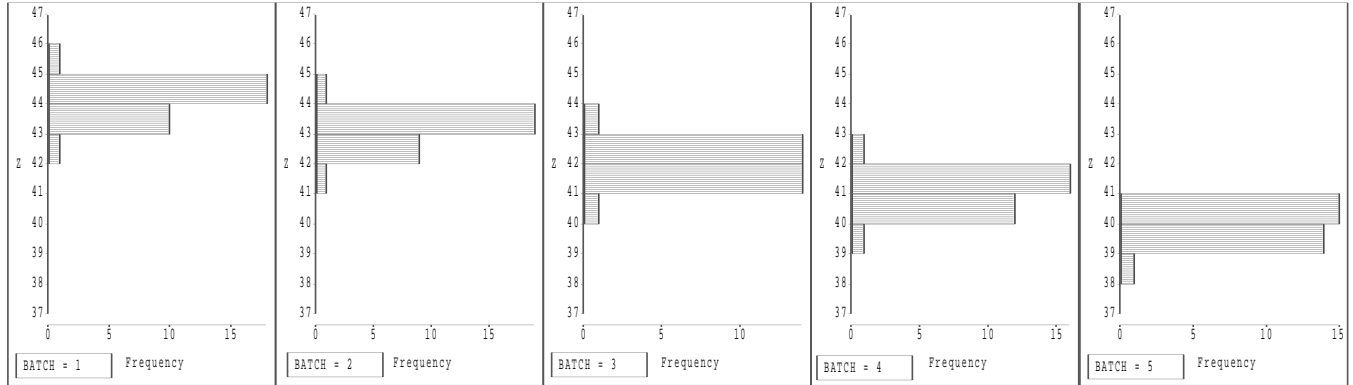
a. Is the process stationary? Why or why not?

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

ANS: *No. The location of the histograms is dropping, from around 44 in batch 1 to around 39.5 in batch 5.* **(10 points)**

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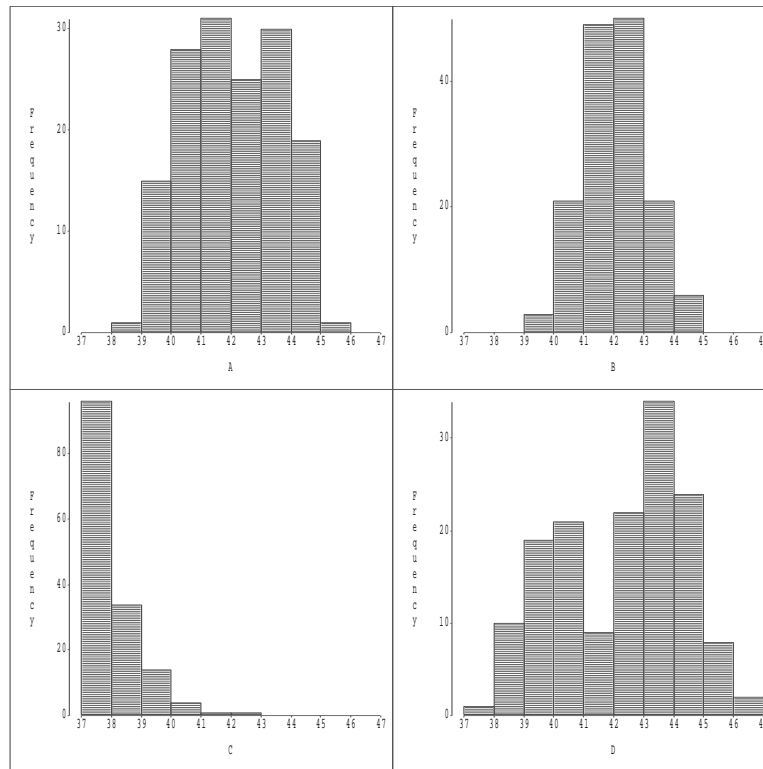

Figure 3: *Four possible Frequency histograms of relative yields, in percent, of production batches of zeolites.*

ANS: *Figure 3A would result.* **(10 points)**

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

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

ANS: *Figure 4 is a stratified plot to assess the pattern of between and within variation. It shows: (i) That Al and W have near-equal within variation, which is substantially larger than Ralph's;* (**10 points**) *(ii) That W and Ralph are on target (i.e., their weight deviations are centered at or near 0), while Al is on the light side, with deviations centered slightly below 0.* (**10 points**)
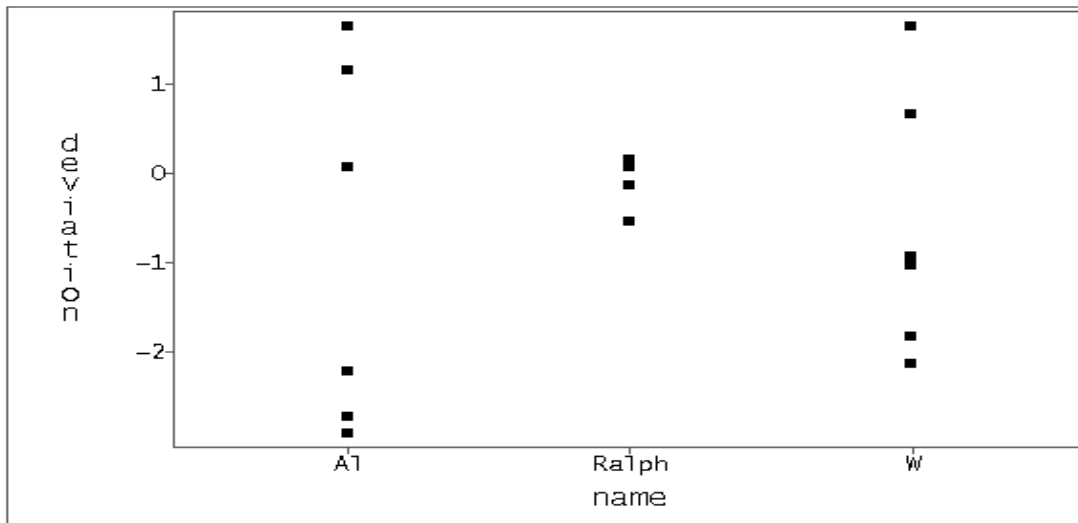

Figure 4: *Stratified plot for S1.3.*