Variation Exploration Lab in a Manufacturing Setting

Diane Evans
Rose-Hulman Institute of Technology, 5500 W. Wabash Ave., Terre Haute, IN 47803

Abstract
In Engineering Statistics, a great deal of attention is paid to variation in a process. In every process, there is variation. Statistical process control charts were developed to separate variation in a process into two components–common cause and special cause. The type that is inherent in a process is called common cause variation; it is there no matter what we do. Variation that drastically affects a process is called special cause variation, and is usually due to just that–special cause. A process is said to be "out of control" if the data plotted on a control chart shows patterns, trends, or unusual jumps. Otherwise, we say the process is "in control." How we sample and group our data to construct a control chart is really dependent on what we want our chart to tell us. For example, control charts can be used to determine if there is variation between our subgroups or even within a subgroup. I have developed a classroom lab in which students work at a company that manufactures Easter peeps. In order to provide answers to their main stakeholders about how "well" the process is running, students must figure out how best to group their data to tell the appropriate "story."

Key Words: variation, process, common cause, special cause

1. Introduction

In Engineering Statistics, a great deal of attention is paid to variation in a process. Since engineering students often collect process data (e.g., time-oriented data), then time series charts or statistical process control charts are often of interest to their studies. Statistical process control charts were developed in the 1920's by Walter Shewhart in order to separate variation in a process into two components-common cause variation and special cause variation. Naturally, there is variation in every process. The type that is inherent in a process is called common cause variation; it is present no matter what we do. For example, every day when I drive to work, the time it takes me to get there varies. Common cause variation in times may be due to stopping at a red light, getting behind a slow car, changes in traffic flow, etc. This type of variation affects my overall drive time, but allows it to stay within "acceptable" time limits from my point of view. Variation that drastically affects a process is called special cause variation, and is usually due to just that–special cause. In my driving time example, special cause variation may be encountering a bad wreck that closes down a lane or a train that blocks my usual path. These incidents will most likely cause a noticeable difference in my driving time, and I can usually identify them on a control chart because they are beyond an "acceptable" limit or "stand out" as significantly different. A process is said to be "out of control" if the data plotted over time is not random; that is, if the data shows patterns, trends, or "unusual" jumps or gaps. Otherwise, we say the process is "in control."

How we sample and group our process data to construct a control chart is really dependent on what we want our chart to tell us. Control charts can be used to determine if the variation in a process from, say, hour to hour, is consistent with the average variation for a process within an hour. A basic principle in constructing a control chart is to choose subgroups in a way that will maximize the probability for the measurement data in each subgroup to be alike and the measurement data between each subgroup to be different.

For example, if our company makes large quantities of chocolate golf balls from molds similar to the ones shown in Figure 1 and our variable of interest is ball weight, then how should we sample and group the balls produced in order to track their weights? Suppose we have a dozen of these molds (with five cavities each) and we use all them to produce 60 golf balls. We could just sample 10 random golf balls from the batch of 60 and average their weights and place that point on an average control chart for 60 batches. We would continue doing this for each batch of 60 balls throughout a day's production. Then we can measure the variability in weight from batch to batch. On the other hand, for the dozen molds, we could sample n = 4 balls from each cavity from a day's production and compare the average weight per cavity. Thus, we would be measuring the variability from cavity to cavity within a mold. There are other sampling plans as well, and the one or ones we end up using should provide the appropriate information needed by various stakeholders of a process to make decisions about the process.

Figure 1. Five-cavity golf ball mold.

The goal of this paper is to introduce a lab from a manufacturing setting that will help students to explore and differentiate between within-subgroup and between subgroup variations within a process. It was designed as an exercise for a Sophomore Level Engineering Statistics course and was given as a one hour in-class lab. The students worked in small groups (3-4 students) on the lab and were allowed to talk amongst themselves and ask me questions. I provided them with data from the process and grouped the data according to three different plans or methods. I have included the questions I posed to them after each plan was executed.

2. Real-world Problem: Three Plans for Subgrouping Manufactured Easter Peeps

There are four machines (A, B, C, and D) that make the identical product: Easter peeps. A certain quality characteristic, say "sponginess factor," is of interest to the main stakeholder, which is the consumer. In this lab, I am considering three different methods or plans of subgrouping the peeps produced by the four machines to be used in control charts. The various machine set ups for allowing the different subgroupings will be displayed in figures for each plan. The students' job as quality control managers is to try to determine which subgrouping plan is best for the various stakeholders (e.g., consumer, boss) in this process.

2.1 Plan A

Select the first subgroup to consist of four observations from machine A, the second subgroup to consist of four observations from machine B, the third subgroup to consist of four observations from machine C, and the fourth subgroup to consist of four observations from machine D. Repeat this grouping method over time. A diagram of this plan is displayed in Figure 2.

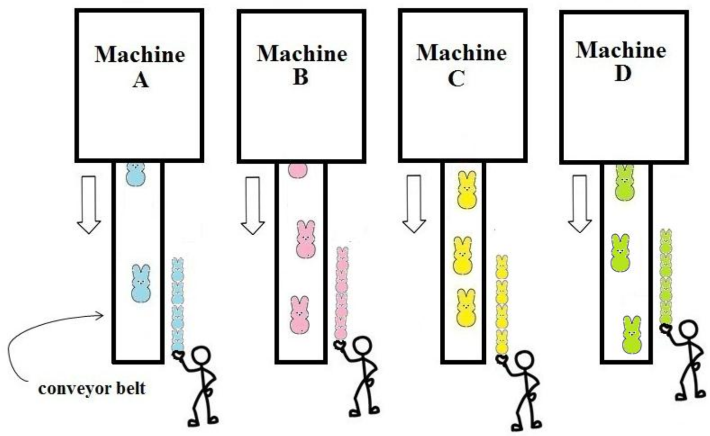

Figure 2. Plan A subgroups by machine. Peeps of the same color are produced per machine. At the end of the line, $n = 4$ peeps per machine are sampled and their average sponginess is recorded.

Two control charts are constructed using 32 subgroups of size four from Machines A, B, C, and D, respectively. The control charts are displayed in Figure 3. For each subgroup of four peeps, the mean sponginess of the four observations is computed as well as the range, where range is simply the maximum minus the minimum value. For example, the first four sponginess observations for Machine A are 7, 8, 9, and 10. Thus, the mean is 8.5 and the range is 3. The values for the mean and range are plotted on their corresponding charts for Machine A. Next, four peeps are drawn from Machine B, and the mean and range are 10.75 and 5, respectively. This process of selecting four random peeps from each machine and computing means and ranges continues for eight cycles. The control charts were constructed using the statistical software Minitab, and the red dots with numbers indicate "out of control" points. Basically, Minitab is alerting us to special cause variation at these points. In order to answer the following questions, it is important to know that the sample means $\bar{x}_i$ for the i^{th} subgroup are plotted over time in the "Sample Mean" chart in Figure 3, the sample ranges R_i for each machine are plotted over time in the "Sample Range" chart of Figure 3, and the red dots note something "unusual" about the process.

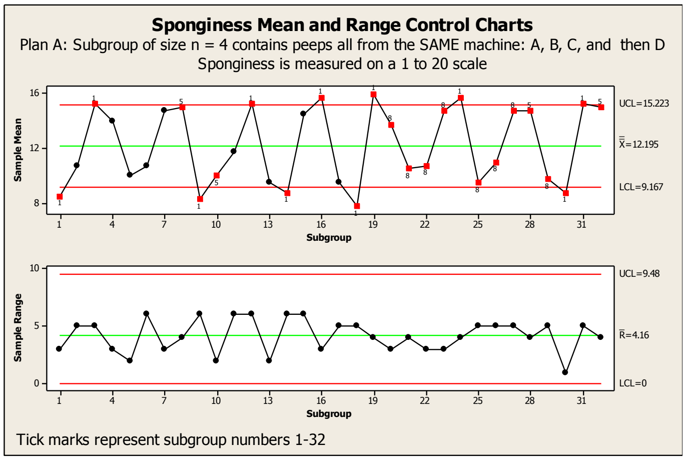

Figure 3. Sample mean ($\bar{X}$) and sample range (R) control charts for subgroups of size n = 4 taken from Machines A, B, C, and then D. The pattern then repeats.

Question 1. Recall that we loosely defined a process to be "out of control" if the data plotted over time is not random; i.e., the data shows patterns, trends, or "unusual" jumps or gaps over time. Otherwise, we say the process is "in control." Would you say that the sample range (R) control chart, shown in Figure 3, is in "statistical control?" Given our subgrouping plan (by machine), what specifically does a given point on the ***R* chart** indicate about variation within a machine?

(A) No, the R chart is not in statistical control. In this subgrouping plan, a point on the R chart tells us that the variation between the machines is not "consistent" or "in control" over time.

(B) No, the R chart is not in statistical control. In this subgrouping plan, a point the R chart tells us that the variation within a machine is not "consistent" or "in control" over time.

(C) Yes, the R chart is in statistical control. In this subgrouping plan, a point on the R chart tells us that the variation between the machines is "consistent" or "in control" over time.

(D) Yes, the R chart is in statistical control. In this subgrouping plan, an "in control" point on the range chart tells us that the variation within a machine is "consistent" or "in control" over time.

Solution: (D) The R chart looks at the variation within a subgroup. Since we are subgrouping by machine, the range is the variation or spread between the four observations from a machine. Although the means are quite different between machines, the spread of observations is "consistent" within a machine.

Answer (C) is incorrect since *a point* on the R chart does not tell us about variation between machines.

Question 2. Would you say that the sample mean ($\bar{X}$) chart, also shown in Figure 3, is in statistical control? Given our subgrouping plan (by machine), what specifically do the points plotted on the $\bar{X}$ chart over time tell us about the variation either within or between machines?

(A) No, the $\bar{X}$ chart is not in statistical control. In this subgrouping plan, the $\bar{X}$ chart tells us that the variation between the machines is not “consistent” or “in control” over time.

(B) No, the $\bar{X}$ chart is not in statistical control. In this subgrouping plan, the $\bar{X}$ chart tells us that the variation within a machine is not “consistent” or “in control” over time.

(C) Yes, the $\bar{X}$ chart is in statistical control. In this subgrouping plan, the $\bar{X}$ chart tells us that the variation between the machines is “consistent” or “in control” over time.

(D) Yes, the $\bar{X}$ chart is in statistical control. In this subgrouping plan, the $\bar{X}$ chart tells us that the variation within a machine is “consistent” or “in control” over time.

Solution: (A) The $\bar{X}$ chart in this plan tells us about variation between machines.

Question 3. What possible explanations can you provide for the way the $\bar{X}$ chart looks in Figure 3? Look again at the control chart and the corresponding data for each point. What do you actually see happening? State your observations in at least two bulleted points and be specific, citing individual machines.

Solution: The first subgroup is from machine A, the second from machine B, the third from machine C, and the fourth from machine D, and this pattern continues.

- Machines A and B have subgroup averages near the lower control limit (LCL), while machines C and D have subgroup averages near the upper control limit (UCL).
- It appears that there are at least two different processes: machines A and B are operating at one average, while machines C and D are operating at a different average.
- There is a correlation between every four plotted points or averages; this is because each machine produces similar values to its previous production.
- The machines are consistent with themselves, but not with each other.

2.2 Plan B

Select exactly one observation from each machine to form a subgroup of four. That is, the first subgroup will contain a peep from Machine A, Machine B, Machine C, and Machine D. Repeat this grouping method over time. A diagram of this plan is displayed in Figure 4.

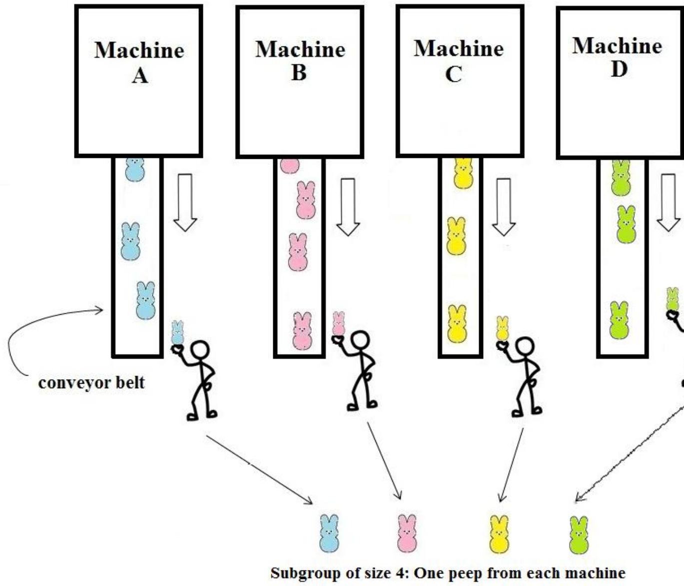

Figure 4. Peeps are produced by machines A, B, C, and D. One peep is sampled from each machine to form a subgroup of size n = 4.

As in Plan A, control charts are constructed using subgroups of size n = 4 from Machines A, B, C, and D. This time one peep per machine forms a subgroup of size four. The control charts are displayed in Figure 5. For each subgroup of four peeps, the mean sponginess of the four observations is computed as well as the range sponginess. The exact same data used to make the control charts in Figure 3 is used to make the control charts in Figure 5; the only difference is the way the subgroups are constructed. Instead of taking four peeps for Machine A, then Machine B, and so on, one peep is selected from each machine to form a subgroup. When determining the mean and range given this new subgrouping method, the control charts take on a very different look. The "story" that these charts give us is totally different than Plan A's. Note, for the first subgroup, Machine A produced a peep with sponginess factor 7, Machine B's peep's sponginess measured 14, Machine C's measured 15, and Machine D's measured 16. The mean and range for this first subgroup are 13 and 9, respectively.

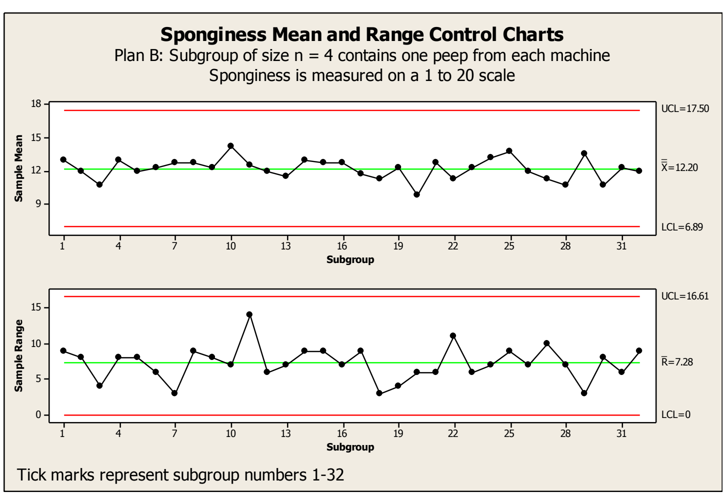

Figure 5. Sample mean ($\bar{X}$) and sample range (R) control charts for subgroups of size n = 4, one peep is taken from each machine per subgroup. The pattern then repeats.

Question 4. Why do you think the average range of the subgroups from Plan B (7.28) is greater than the average range of the subgroups from Plan A (4.16)? (The mean range $\bar{R}$ is displayed as the center line of the R charts.) What does this say in the context of the two subgrouping plans?

Solution: In Plan B, one observation is taken from each machine to determine the range for each subgroup. As we saw from Plan A, the observations from Machine A and Machine B are "low," while the observations from Machine C and Machine D are "high." For example, in the first subgroup, Machine A produced a peep with sponginess factor 9, Machine B's peep's sponginess measured 14, Machine C's measured 15, and Machine D's measured 16. The range for this subgroup is high (relative to ranges from Plan A) because of the large "spread" of the observations produced from these four different machines.

Grouping the data from within a machine versus between four machines produces a different value of the range, and thus a different standard deviation of the process. Which is the better estimate or "more correct" of the two? This answer depends on what you are using your charts to determine and definition of variance you are using: between subgroup variation or within subgroup variation.

Question 5. True or **False**. Since the sample range R chart is in control for Plan B, this means that the variation within a subgroup across the four machines is basically about the same from subgroup to subgroup. Although the sponginess factor ranges between machines may be "large," it is consistent.

Solution: True!

Question 6. Although the sample mean $\bar{X}$ control chart appears to be in statistical control, it is very close to breaking an "out of control" rule that is signaled when there are "15 points in a row within one standard deviation of the center line." What explanations do you have for why the $\bar{X}$ chart looks this way? That is, why are the subgroup means hugging the center line for the majority of the plot? (Hint: The upper and lower control limits on the $\bar{X}$ chart are obtained by the measure of dispersion per subgroup, which in term is calculated using the average range $\bar{R}$.)

Solution: Machines A and B are producing peeps with low sponginess factors, while machines C and D are producing peeps with high sponginess factors. When we sample peeps from machines A and B, we get low values; when we sample from Machines C and D, we get high results. Averaging these four observations produces an overall average that is consistent from subgroup to subgroup. In addition, the process standard deviation for Plan B, computed using the average sample range $\bar{R}$ for Plan B, is larger than the process standard deviation for Plan A. Thus, the upper and lower control limits are "further away from" the center line on the sample mean $\bar{X}$ chart for Plan A. Thus, averages seem to plot closer to the center line because of the amplified upper and lower control chart limits.

2.3 Plan C

Select each subgroup to consist of four peeps from the blended stream. Each machine's conveyor belt merges the peeps onto a larger conveyor belt creating a blended stream of peeps. We will assume the peeps are coming down the conveyor belt so fast that the worker will disregard colors and just group n = 4 on the belt in same time vicinity. Assuming the line worker disregards colors and does not record them in his data report, it is impossible to know or go back and track which machine manufactured a given peep in a subgroup. The diagram in Figure 6 conveys this idea.

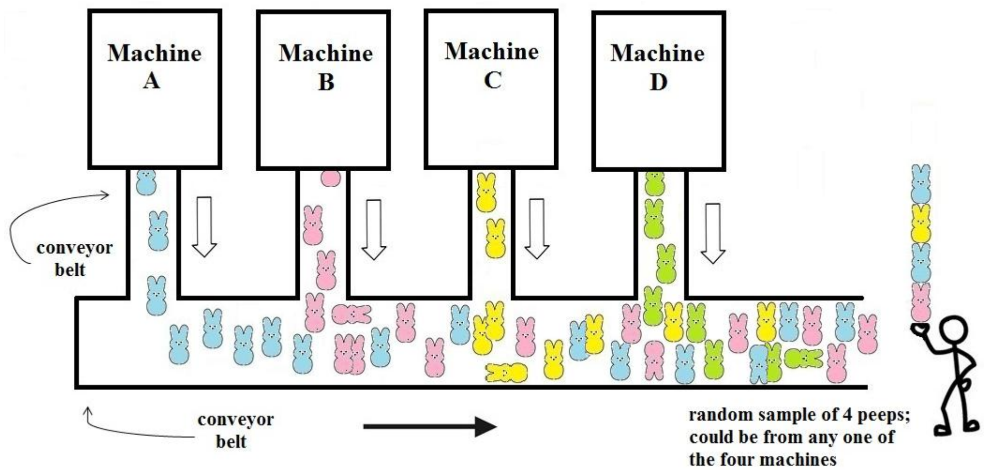

Figure 6. Peeps are produced by machines A, B, C, and D and blend together on the final conveyor belt. The subgroup of n = 4 peeps at the end of the conveyor comes from a "blended stream."

As in both previous plans, control charts are constructed using subgroups of size n = 4 from Machines A, B, C, and D. Since the peeps are dispersed from the machines onto the larger conveyor belt without a concern for order, the worker at the end of the conveyor belt samples four peeps without regard to which machines they came from. The control charts for this plan are displayed in Figure 7. For each subgroup of four peeps, the mean sponginess of the four observations is computed as well as the range sponginess. The same data used to make the control charts in Figures 3 and 5 is again used, but this time the output from each machine is randomly assigned to a subgroup. For example, the first subgroup may contain peeps from only Machines A and B, or just A alone. On the other hand, the next subgroup may contain peeps from all four machines. Again, the problem is that when we subgroup peeps, we cannot tell which machines produced the peeps.

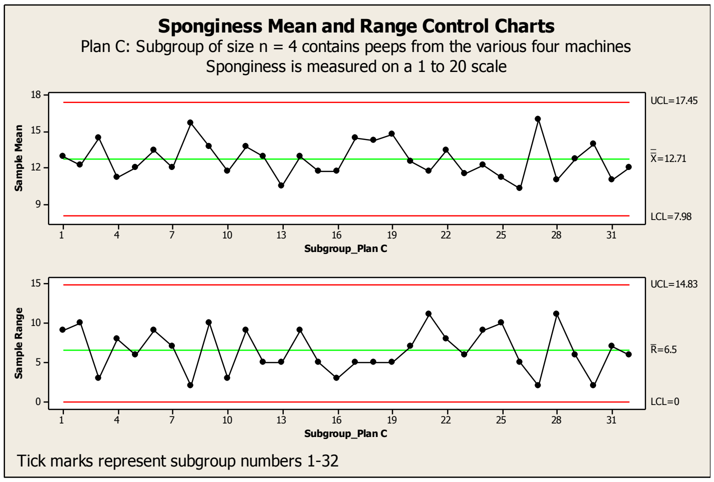

Figure 7. Sample mean ($\bar{X}$) and sample range (R) control charts for subgroups of size n = 4 taken from Machines A-D, although which particular machines is unknown. Peeps produced by the machines are blended together, and so the subgroups come from a "blended stream."

Question 7. Both the sample mean $\bar{X}$ and range R control charts are in statistical control. Do you see a problem with combining all of the data into one stream and selecting subgroups from that stream? What can those charts tell us about our process?

Solution: Yes, we do not know which machines are producing the observations that make up the subgroups. There is always the chance that we could be sampling from just one machine if the stream is not well "mixed." So, it is hard for us to identify if there are any problems between the machines. If the process does go out of control, we do not know if it is due one machine, several machines, or some other problem in the system.

These charts do tell us that our process is in control, both in terms of spread and center. As we can see from Plan A, the machines are producing peeps with different sponginess factors. Although the process is in control, it is likely not going to meet specifications set by the customer because of the variation between machines. The charts above are "masking" the problem of different sponginess levels between machines.

Question 8. The average range $\bar{R}$ of the subgroups for Plan C is 6.5, while it is 4.16 and 7.28 for Plans A and B, respectively. Does this make sense?

Solution: Yes.

Plan A: $\bar{R}$ is the average of the ranges *within a machine*. The ranges are the smallest for observations taken from only one machine.

Plan B: $\bar{R}$ is the average of the ranges *between the four machines*. Since the machines are producing peeps with varying sponginess factors (both low and high), it is not surprising that the range is largest in this plan.

Plan C: Since $\bar{R}$ is a blend of observations from one or more machines, we would expect $\bar{R}$ to be between the values obtained from Plan A and Plan B.

Question 9. The overall mean for samples from Plans A, B, and C are quite similar. Why?

Solution: Plans A and B use exactly the same data, except "grouped" differently – that is why the center lines are the same on their $\bar{X}$ control charts, which is the mean of the means. For Plan C, the overall mean should be similar, but slightly different because we are not using exactly the same observations. More (or less) observations from one of the machines would cause a higher or lower overall average.

3. Comparison of All Three Plans

Suppose that consumers really do care about the sponginess of peeps. Perhaps kids from all over the world will only eat them if they are "squishy enough" (but not too squishy for parents to clean up the mess). In other words, suppose that the sponginess factor of peeps is very important.

Question 10. Assuming your major stakeholder for peeps is the consumer, which subgrouping plan of the three presented is the "best" in your opinion? Why? How are you defining best?

Solution: Let us define "best" as the plan that actually tells us something useful about the process, and in turn, helps us produce a better product for the consumer.

If we are really interested in telling whether or not the machines are producing the "same" product with respect to the sponginess factor, then Plan A is best. Plan A shows the discrepancy between machines in the sample mean $\bar{X}$ chart. After viewing Plan A's chart, we can make adjustments to the machines to "center" their production with respect to sponginess. A peep that is either too spongy or not spongy enough could wreak havoc with a child consumer's Easter pleasure. In order to please our "tiny" consumers, we would like to produce a consistent product. Every box of peeps should have the same sponginess; customers expect a certain level of consistency when they continually buy the same product.

Plan B is "too good to be true" and tells us very little about the process since we are averaging observations between all four machines that result in subgroup averages that are near the center line. In fact, Plan B could be considered the worst plan for the consumer because it does not alert us (the producer) to change our production pattern (unless a skilled quality expert is studying the average range). If we do not make changes to the machines, then the sponginess factor of these peeps will vary widely.

Plan C can tell us if the material being produced is "consistent," but it does not give us information about the individual machines' productions. We may or may not (depending on the results produced by the subgrouping) take a better look at the differences in machines' productions. So, again, the consumer may "lose out" if we do not correct our production process.

Question 11. Assuming that your boss is the only stakeholder that you concerned about, what would be the worst plans' control charts to show your boss? Hint: Although these charts may get you a bonus in salary in the short term, they may get you fired in the long term.

Solution: Again, Plan B is "too good to be true" and tells us very little about the process since we are averaging observations between all four machines that result in subgroup averages that are near the center line. In fact, Plan B is the worst because your boss will think the process is in control (because the values are so tightly clinging to the center line), when in fact, your control chart is masking any potential problem you may have between the machines.

Question 12. If the boss gave you free reign to develop your own sampling plan, what would you do? Think as far "out of the box" as you would like.

Solution: One popular suggestion is to consider Individual (*I*) and Moving Range (*MR*) charts for each machine. Students come up with many additional ideas for this.

4. Conclusions

In this lab, students must apply their knowledge of statistics to process (rather than population) to further explore the concept of variation. On top of that, they can apply their statistical background to a real-world problem situation.

The lab's questions, specifically Questions 10-12, require the students to view the subgrouping in the manufacturing process from not only the consumer's viewpoint, but also the viewpoint of the boss of the company. *USA Today* is a purveyor of charts and graphs, and even occasional readers are keenly aware of the daily Snapshots that summarize data using fun and interesting graphics. Experienced statisticians and other professionals know how to get a certain meaning across by choosing the appropriate graphic – it is a desirable skill. Students must be careful to accurately present their data, while optimizing its effect on its audience or stakeholders. What may be an appropriate graphic to show "in house" for improvement of a process may not be the right one to show your board of directors.

In this lab, students must associate real-world management objectives with statistical ideas. They must question the "status quo," which in many manufacturing settings is simply letting output from different machines mesh together in a stream and sample from it (as in Plan C). Personally, I have heard dozens of statistical consulting situations where clients did not understand the problems created when output from different machines all funneled the same product into one line. Also, students are encouraged to experiment with the actual data (given to them in a statistical software package) and rearrange it in various other ways to obtain other views of the process. Often several students decide to build control charts for individual data points instead of means, which bring up new questions as to why sample means were used in the first place.

5. Further Considerations or Extensions

I plan to extend this lab with an actual simulation of the process (with actual Easter peeps, though the factor of interest may be something more easily measured, such as length), more detailed instructions and data (without preparing it beforehand already grouped by certain plans), and additional follow up and research questions to include in my Quality Control and Six Sigma courses. For example, here are some extensions of the assignment:

- Choosing suppliers for materials (e.g., sugar, food coloring)
  - → Are they ISO certified?
  - → Compare costs, quality, reliability
  - → Acceptance sampling (i.e., plan for sampling from supplier and rejecting bad lots)
- Determining machine set-up
  - → Manual versus automatic operation
  - → Practical limitations of flow through the entire system
- Determining variables to measure (e.g., weight, sponginess, color quality)
  - → How to measure? Is the necessary equipment available?
  - → At what point in the system should measurements occur?
  - → Is the variable being measured discrete (count data) versus continuous (measurement data)?
- Sampling
  - → Frequency and size (e.g., four every 15 minutes versus eight every 30 minutes)
  - → Subgrouping (as discussed in this lab)
  - → Sampling interval (e.g., sample every 5th item or take a group all at the same time)
- Handling of defects
  - → Rebuild (at what cost)
  - → Scrap (at what cost)
- Applying quality control to the system for improvement
  - → What type of control chart to use
  - → Perform a Gage R & R (repeatability and reproducibility) on the machines and workers
  - → Perform a capability analysis

Building an entire system to produce the peeps from start to finish will require them to draw on their engineering background and remove themselves momentarily from just statistical thinking. A process map would be an ideal exercise in conjunction with these extensions. Questions about when to sample and how often to sample will require brainstorming as a group, and there is no one “correct” answer. In order to understand Gage R & R, they will need to understand how two people could measure the same entity and get two different values. More surprising than that is how one person can repeatedly measure one object and get different values. Through experimentation with 10 peeps, two operators, and two calipers, it is easy to show that when randomizing the order of peeps to be measured without operator knowledge, the measurements will rarely matchup between operators, or even two measurements by the same operator.

Finally, this lab can easily be adapted for a one or two hour workshop activity in Six Sigma. As I have just started teaching a class in Six Sigma at my home institution, I plan to extend this activity throughout the course as we move through the many topics as part of the DMAIC process.