Ryerson University
Faculty of Engineering & Architectural Science
Department of Mechanical and Industrial Engineering

Program: Mechanical Engineering/Industrial Engineering

Wheel Liner and Air Deflector

Quality Assurance

by Abdul Sami Ali, Kevin Fong, Suhasee Patel, and Tahsin Rezaul

Abstract

The manufacturing process of a Wheel Liner and an Air deflector will be analyzed using data samples taken from 4 different scenarios, 1 base scenario with no improvements and 3 scenarios with varying improvements. The current manufacturing process can be optimized and improved upon as there are health and safety concerns with the current process, specifically the manufacturing process of the Wheel Liner. Using visual (various charts) and statistical analysis (ANOVA and hypothesis testing) scenario 4 which had both technical and operational improvements was found as the best result. Using the improvements given in scenario 4, further research operational and technical improvements is needed , specifically the type of arm tooling robot that should be implemented as there are different types that bring different advantages to the process.

Introduction

The subject matter that is being analyzed and discussed is the manufacturing process of a Wheel Liner and an Air Deflector. The current process has some flaws including a poor management system due to operators doing work that could be done by a computer and a hazardous process when producing the Wheel Liner. To improve the process, the company is testing various scenarios with technical (optimization-based and an automated quality control system) and operational improvements (arm tooling robots and an automated monitoring system). The design problem that the rest of the paper will be looking to solve is which scenario optimizes the manufacturing process. A detailed visual and statistical analysis will be done in the body of the report which will result in a discussion of the best scenario and possible alternative solutions. A conclusion will follow, which will state and explain the final recommendations that the group suggests will best optimize the process. The appendix at the end of the report has all the charts and hypothesis testing for all the settings for the throughput and non conforming units data.

Approach

The purpose of this report is to identify which scenario best optimizes the manufacturing process of a Wheel Liner and an Air Deflector. Data was provided on 2 quality characteristics (throughput and number of non conforming units) for 4 different scenarios of a manufacturing cell; base manufacturing cell with no improvements, manufacturing cell with operational improvements, manufacturing cell with technical improvements and manufacturing cell with technical and operational improvements. Only those improvements were made for each scenario and the data provided was for an average operator are key assumptions that were made. To identify which scenario is best for improving the process, visual and statistical analysis was performed in the form of control charts (X-bar, R and P), histograms, time studies and box plots for the former and ANOVA and hypothesis testing for the latter.

Main Body

An optimization based system analyzes issues and activities that may arise in a workplace in the form of discrete or continuous data (Kacem, 2017), and answers questions such as:

- What tasks must be executed to satisfy the given demand (batching/lot-sizing)?
- How should the given resources be utilized (task-resource assignment)?
- In what order are batches/lots processed (sequencing and/or timing)?'
  (Chenn & Grossmann, 2016)

By analyzing and comparing various charts and their characteristics, conclusions are drawn to produce an optimized based system. Production scheduling is used to be efficient in a work process by adequately balancing customer needs with the resources that are available whilst operating in a cost-effective manner (PlanetTogether, 2018).Using an optimization based system approach to generate production schedules would ensure the production schedule is fully optimized.

Visual analysis

X-Bar Chart: A type of control chart that is used to monitor the means of constant samples n. This type of chart is generally used for measurements on a continuous scale. It typically tells us how the mean changes over time.

R-Bar Chart: A type of control chart that shows the range of various subgroups and how they change overtime. This type of chart can help when it comes to monitoring the effects of process improvement theories and how they affect each operation.

Time Series: A time series chart, shows data points at successive time intervals and is used to see how the data changes over a period of time. As the variation of the data decreases, the line of data points will look more clustered and close together.

Box Charts: They provide a quick visual representation of the data distribution through their quartiles and where the range is in accordance to the quartiles (DatavizCatalogue, n.d.). The yellow box represents the amount of data between the first quartile and median whereas the green box represents the amount of data between the median and third quartile. The lines extending parallel to the boxes are known as the whiskers which indicate the variability outside the upper and lower quartiles. For the data given for NU, it is desirable to be closer to zero as that represents less products being non-conforming.

Histogram: Provides a visual representation of the frequency of the data per lot analyzed (DatavizCatalogue, n.d.). It is easy to understand at which intervals the data is most frequent.
Given the data for NU, having frequency at lower numbers is desirable as that represents that there are less non-conforming units compared to the rest of the settings. Given the data for TP, having frequency at higher numbers is desirable as that represents the amount of work being done in a certain period of time.

P Chart:

In statistical quality control, the p-chart is a type of control chart used to monitor the proportion of nonconforming units in a sample, where the sample proportion nonconforming is defined as the ratio of the number of nonconforming units to the sample size, n. Having significant values within the control limits means a lower proportion of non conforming units.

Implementing an automated quality control system ensures continuous quality improvement of a process. This system consists of an understanding of analytical errors among many other data to determine if the analytical run is within control or out of control by the quality control standards. These standards are defined to ensure the process returns an acceptable control (Badrick, 2008).

Quality control dashboards are applications for quality reporting and population management. By providing various benchmarks, data can be analyzed and optimized to maintain efficiency and quality (Olsha-Yehiav, Einbinder, Jung, et al., 2006).

Transitioning from a quality control dashboard to a quality control system ensures that the process being carried out is to be done within control rather than adjusting the process to ensure it comes to be within control.

Throughput analysis

Quality control charts are a great tool to monitor whether processes are in statistical control. They allow the viewer to interrupt variation, find and correct problems and predict expected ranges and outcomes. They are also often used to analyze patterns of process variation due to special causes.

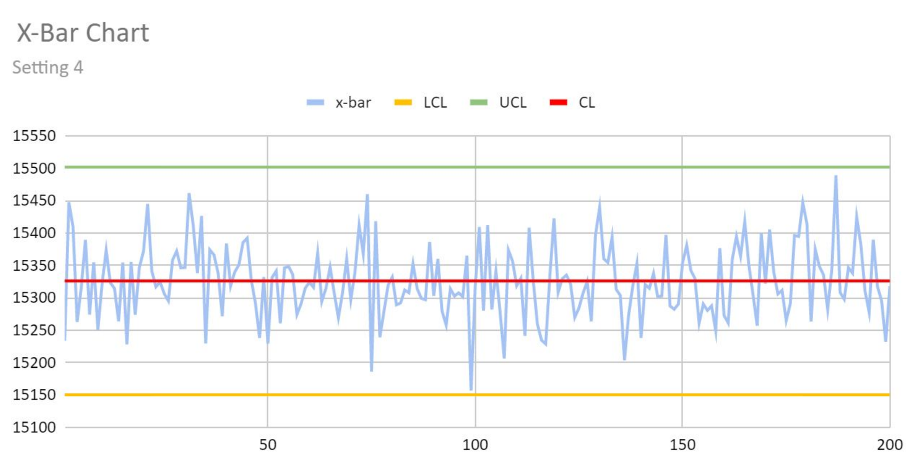

Figure 1.4: X-bar Chart, Setting 4

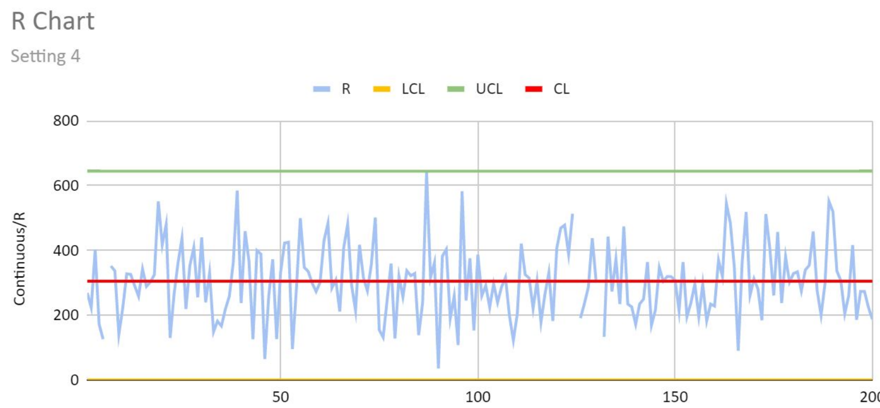

Figure 2.4: R Chart, Setting 4

In order to sufficiently analyze all four settings, we took a look at the R-charts first and determined if there were any points that crossed the upper and lower control limits. Settings 3 and 4 had no significant data points out of control meaning that the mean and the variation of the processes were in control. In order to determine the best setting for the company to maximize the throughput we made sure to ensure that all of the process points were in control. Taking a look at all four of the settings (can be found in the appendix), setting four had created the ideal x bar chart because all of the process points were in control meaning that there is no change in the process required. There is an even distribution of data points within the x bar chart for setting 4 meaning that process is under statistical control. The other settings had many data points that were not within the control limits. This means that the other settings go beyond limits, showing us that process may not have had enough time to be completed. This setting proves that implementing both technical and operational changes to the process maximizes the throughput.

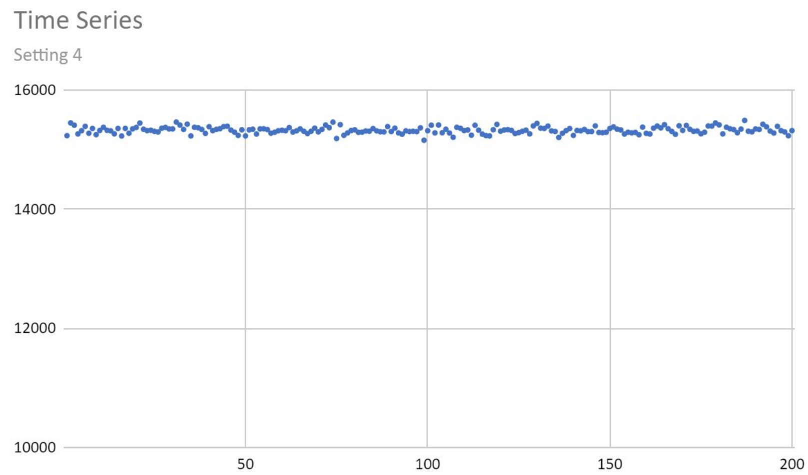

Figure 3.4: Time Series, Setting 4

Between the 4 time series charts, setting 4 created the best time series chart as it has the lowest variation between all the points because the data points were the most clustered. As seen in the graph above, the throughput is the highest in setting 4 and the points are very clustered around 15500 meaning the variability is low.

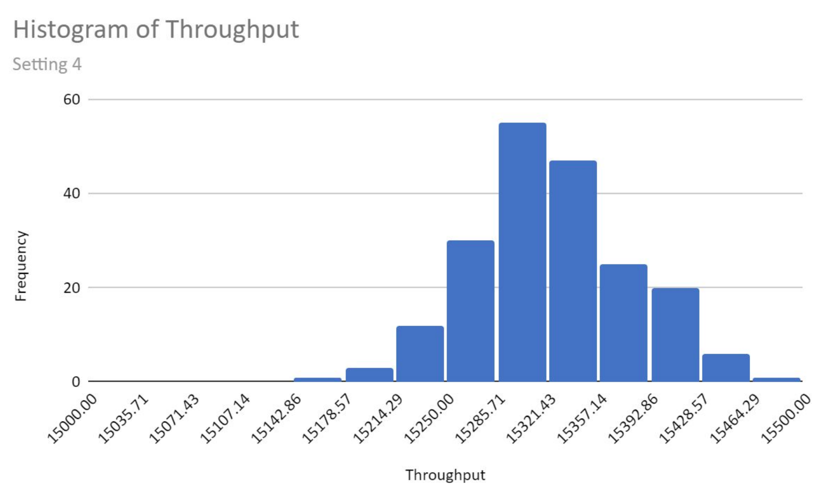

Figure 4.4: Histogram, Setting 4

After analyzing the 4 histograms for throughput, the histogram for setting 4 shows the data being frequent at higher values. This is desirable for throughput as throughput represents the amount of work being done in a certain period of time. The higher the number, the more work being done per set time.Observing the data provided, setting 4 yields the most work being done per set time ranging from 15142.86 to 15500.00, followed by setting 3 ranging from 13785.71 to 14214.29.

Number of Non-Conforming Units analysis

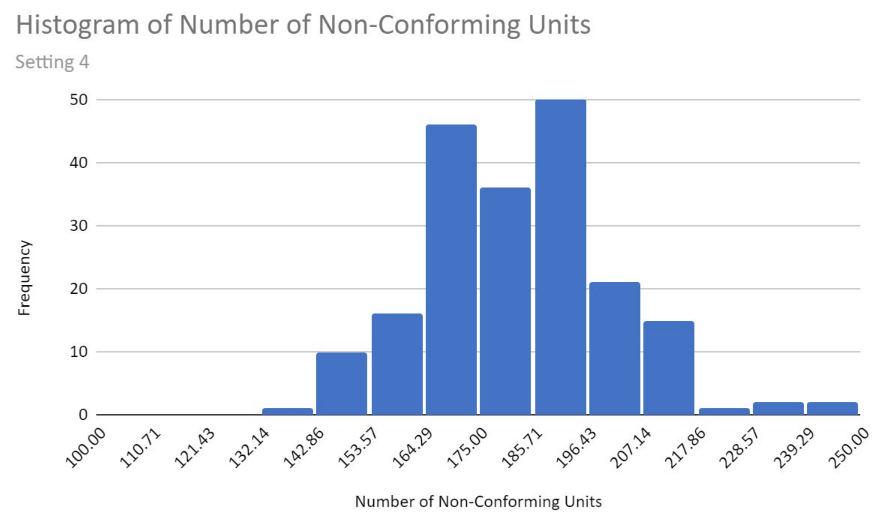

Figure 5.4: Histogram, Setting 4

After analyzing the 4 histograms for number of non-conforming units, the histogram for setting 4 yields less non-conforming units compared to the other settings. Having frequency of non-conforming units at lower numbers is desirable as that represents that there are less non-conforming units compared to the rest of the settings. Observing the data provided, setting 4 yields the lowest number of non-conforming units with its mass ranging from 132.14 to 250.00, followed by setting 3 with non-conforming units ranging from 164.29 to 335.71. Analyzing the histogram depicts that setting 4 yields the most desirable set of data out of the 4 settings.

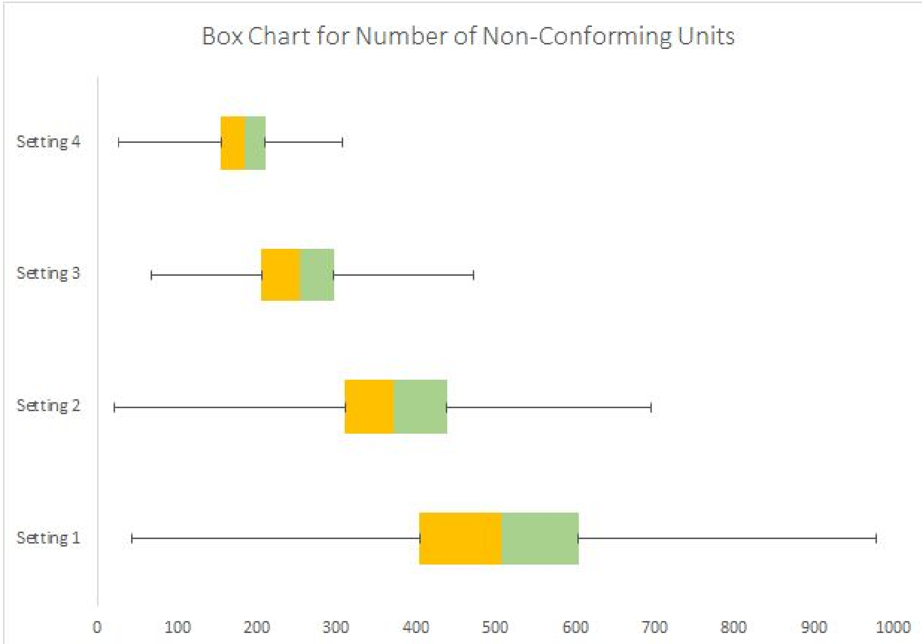

Figure 6.5: Box Plot, All Settings

After analyzing the 4 box charts for number of non-conforming unit, the box chart for setting 4 yields smallest range of data with most of the data points being closer to zero, compared to the other settings. This is desirable for box plots as that represents less products being non-conforming. Though setting 2 has its whisker closer to zero than that of setting 4, the box (representing the mass data) being of a higher quantity suggests that it may be an outlier. Understanding such, it can be concluded that setting 4 is the most desirable for NU.

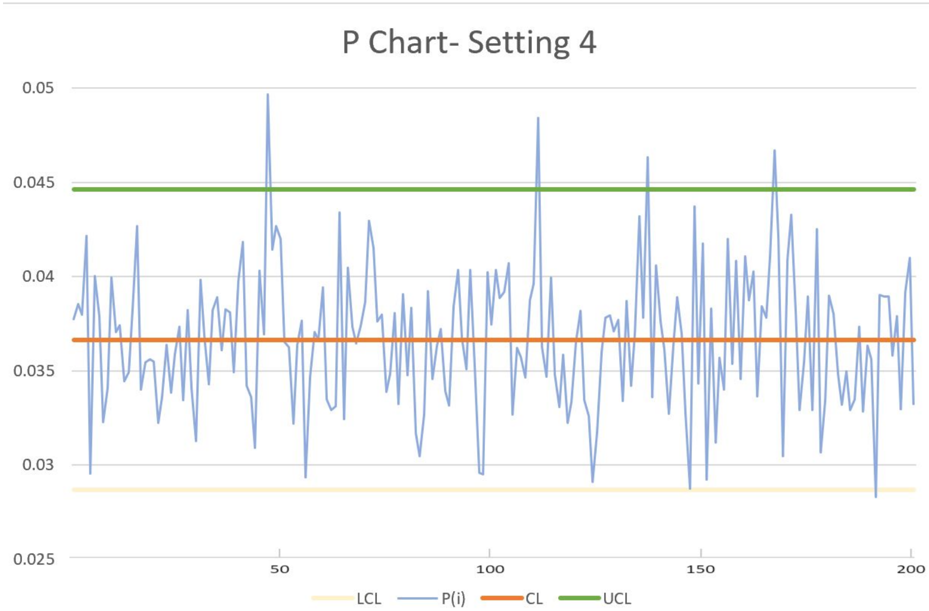

Figure 7.4: P Chart, Setting 4

After analyzing 4 settings for the p chart, settings 4 has no significant data points out of control of the control limits meaning less fractional non conformity. Setting 4 creates the most ideal p chart because almost all of the process points were in control. The other settings had many data points outside the control limits. This means that the other settings go beyond limits, showing us that many of the units produced did not conform to the standards. This setting proves that implementing both technical and operational changes to the process minimizes the number of fractional non conforming units.

Statistical analysis

From the ANOVA for TP and NU we can reject the null hypothesis that the TP and NU are not affected by the operational and technical improvements because the F value is greater than the F critical value meaning that both TP and NU are affected by the operational and technical improvements.

Hypothesis Testing Summary

Table 1. Summary of all Hypothesis Testing conducted for Throughput

Settings	$\lvert(t_o)\rvert$	$t_{\frac{\alpha}{2},\, n_1+n_2-2}$	Reject null hypothesis or Fail to reject null hypothesis
1-2	155.2407	1.96	Reject null hypothesis
1-3	265.7627	1.96	Reject null hypothesis
1-4	461.389	1.96	Reject null hypothesis
2-3	124.542	1.96	Reject null hypothesis
2-4	340.249	1.96	Reject null hypothesis
3-4	203.708	1.96	Reject null hypothesis

*refer to appendix for full analysis

When comparing each of the settings to one another our team used two sample hypothesis testing because the data was normally distributed so we were able to test the null hypothesis for a single population using a two sample pooled t test. Taking a look at the results from the test statistic we can conclude that when taking a look at the various settings the t test was rejected. The null hypothesis can be rejected at the significance level of 0.05 due to the sufficient evidence within the sample that favours the alternative hypothesis. Meaning that for all changes made, technical, operational or both there was an improvement within the settings because the overall mean was different for each. It can be observed that setting 4 had the highest mean proving that with both technical and operational improvements, this setting was able to produce the most amount of units within the time it takes to complete the operation.

Recommended improvements

Through the visual and statistical analysis, setting 4 would yield the best results to improve the manufacturing process. As stated in the project summary the technical improvements would be an optimization-based and an automated quality control system and the operational improvements would be arm tooling robots and an automated monitoring system.

The monitoring system would minimize downtime and allow workers to do predictive maintenance which will optimize the maintenance costs (Stavropoulos et al., 2013). Downtime is costly as no manufacturing is done while the machines don't work and customers may lose trust as your manufacturing is unreliable. By monitoring your process, you will have access to data 24/7 and as soon as something breaks, you can immediately address the issue and try to fix it(Wasmund, 2019). It will also allow you to do predictive maintenance as you will be able to track important variables like temperature and vibration levels, therefore you can prevent accidents or breakdowns before they happen. Overall maintenance will happen proactively instead of reactively which will save time and money in the future.

The end of arm tooling robots (EOAT) would speed up the manufacturing process while making it safer for the operator as they wouldn't have to do the hazardous process of placing the fabrics into the cavity side of the mold. Further recommendations for the EOAT are deciding which gripper and sensor the company should get as each type has varying advantages and disadvantages. For example there are 3 different types of grippers it can come with: mechanical, vacuum or adhesive (What is end of arm tooling for robots?, 2020). Mechanical uses mechanical fingers that use a pincer motion to pick things up, vacuum uses a suction cup to grip things and adhesive grips objects by sticking to the object. When picking up fabrics, adhesive is mostly used because fabrics are lightweight objects however the reliability of the adhesive goes down after each operation because it loses the adhesive quality. Testing of the breakpoint of the adhesive and whether it is worth it from a cost perspective is recommended for a next step. EOATs also come equipped with 2 main types of sensors: collision and force sensors. Collision sensors shut the arm done and reset the arm back to its original position if it detects a collision incoming while the force sensors provide force and torque measurements for the arm. Since this arm will just be picking up fabrics, it is recommended that only the collision sensor is needed

Conclusion

In this project, the manufacturing process of a wheel line and air deflector were analyzed using the data samples from 3 different scenarios. After performing statistical analysis and making visual charts, setting 4 was determined to be the best scenario which had both technical and operational improvements. As per the ANOVA analysis for TP and NU, the null hypothesis was rejected because the F value was greater than F. The test statistic helps us conclude that t test was rejected for various settings. Moreover , the Z-test was also rejected for the discrete data set. Technical improvements would be an automated quality control system and the operational improvement would be an arm tooling robot and an automated monitoring system. The monitoring system would help track important variables and allow us to prevent any accidents. Using an adhesive gripper is recommended for the end of arm tooling robot. In addition, having collision sensors would help shutting the arm down if a collision is detected. Overall, this report endorses the fact that applying technical and operational improvements optimizes the whole process to a significant level.

References

Badrick T. (2008). The quality control system. *The Clinical biochemist. Reviews*, *29 Suppl 1*(Suppl 1), S67–S70.

Chen, Q., & Grossmann, I. E. (2016). Recent developments and challenges in optimization-based process synthesis. *1 Center for Advanced Process Decision Making,* 1-5.

DatavizCatalogue. (n.d.). Box and Whisker Plot. Retrieved from https://datavizcatalogue.com/methods/box_plot.html

DatavizCatalogue. (n.d.). Histogram. Retrieved from https://datavizcatalogue.com/methods/box_plot.html

Kacem, I., Dr. (2017). Applications of Optimization Methods to Industrial Engineering Problems. *Computers & Industrial Engineering.* 112. 467-468. 10.1016/j.cie.2017.08.034.

Olsha-Yehiav, M., Einbinder, J. S., Jung, E., Linder, J. A., Greim, J., Li, Q., Schnipper, J. L., & Middleton, B. (2006). Quality Dashboards: technical and architectural considerations of an actionable reporting tool for population management. *AMIA ... Annual Symposium proceedings. AMIA Symposium*, *2006*, 1052.

PlanetTogether. (2018). 4 Components of Production Scheduling in Manufacturing. Retrieved from https://www.planettogether.com/blog/4-components-of-production-scheduling-in-manufacturing#:~:text=Production scheduling is the allocation,to create goods and services.&text=The goal of a production,in a cost-effective manner

Stavropoulos, P., Chantzis, D., Doukas, C., Papacharalampopoulos, A., & Chryssolouris, G. (2013). Monitoring and Control of Manufacturing Processes: A Review. Procedia CIRP, 8, 421–425. https://doi.org/10.1016/j.procir.2013.06.127

Wasmund,R (May 28,2019). Why is Remote System Monitoring Important? Concept System Inc. https://conceptsystemsinc.com/best-remote-monitoring-system-for-industrial-automation/

What is end of arm tooling for robots? (2020, April 20). OnRobot. https://onrobot.com/en/blog/what-is-end-of-arm-tooling-for-robots

Appendix

Throughput

X-Bar Chart

Figure 1.1: X-bar Chart, Setting 1

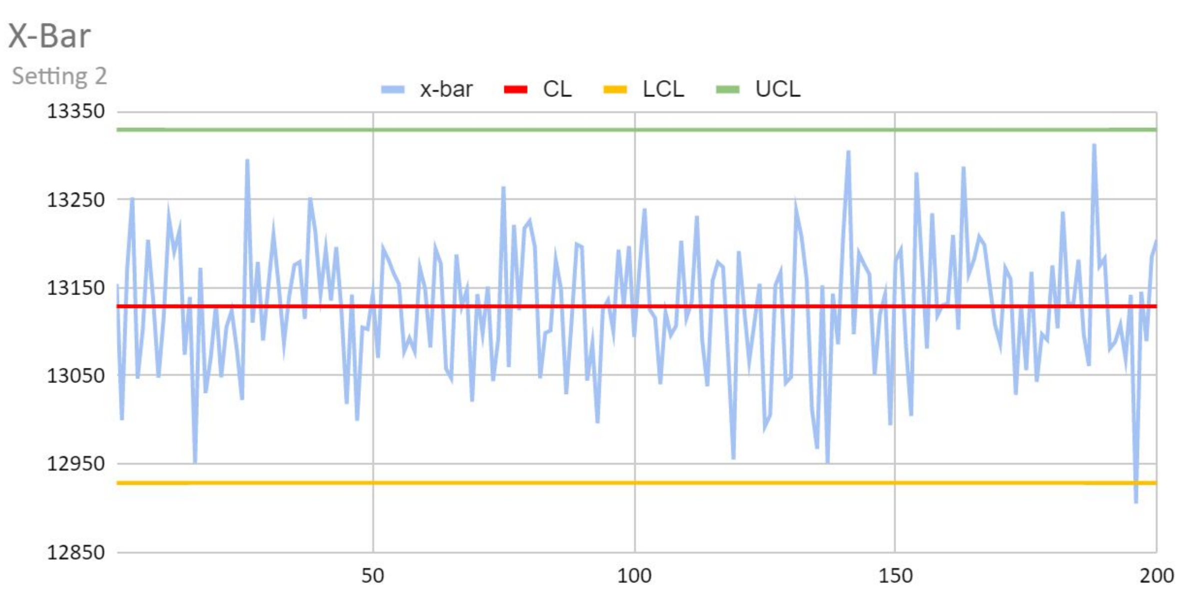

Figure 1.2: X-bar Chart, Setting 2

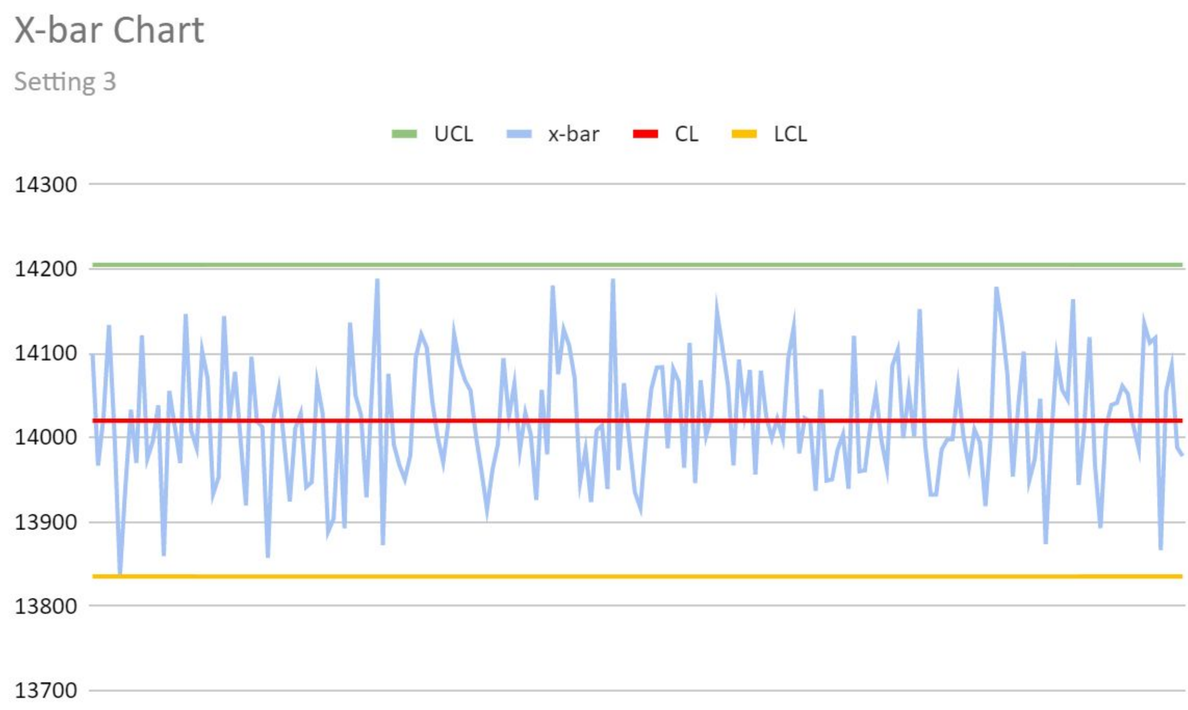

Figure 1.3: X-bar Chart, Setting 3

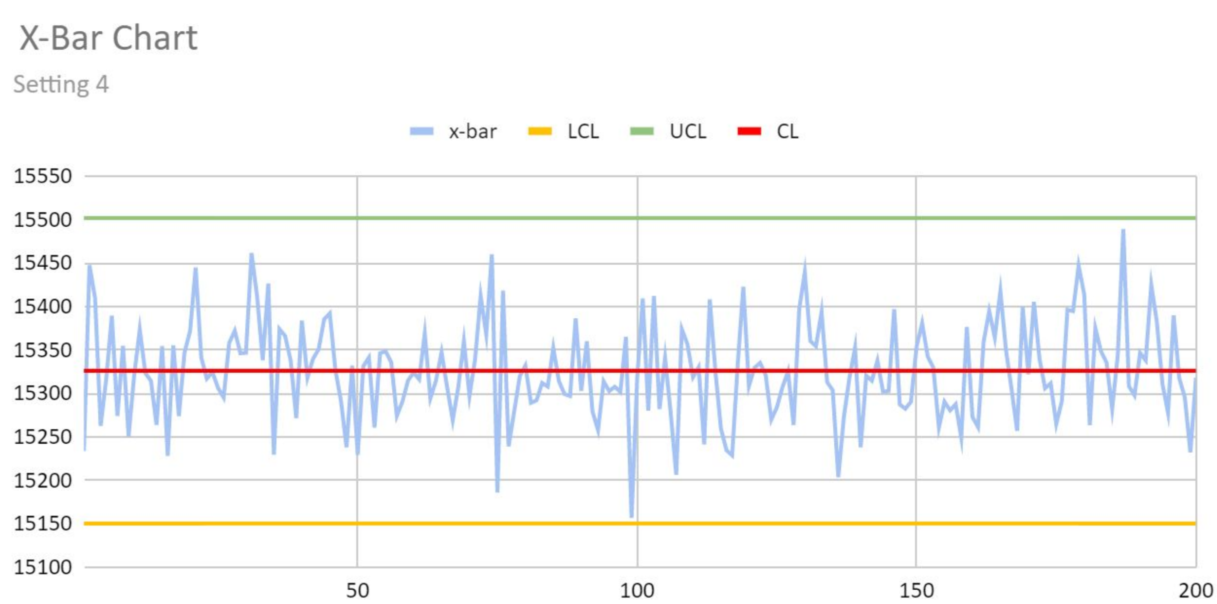

Figure 1.4: X-bar Chart, Setting 4

<u>R Chart</u>

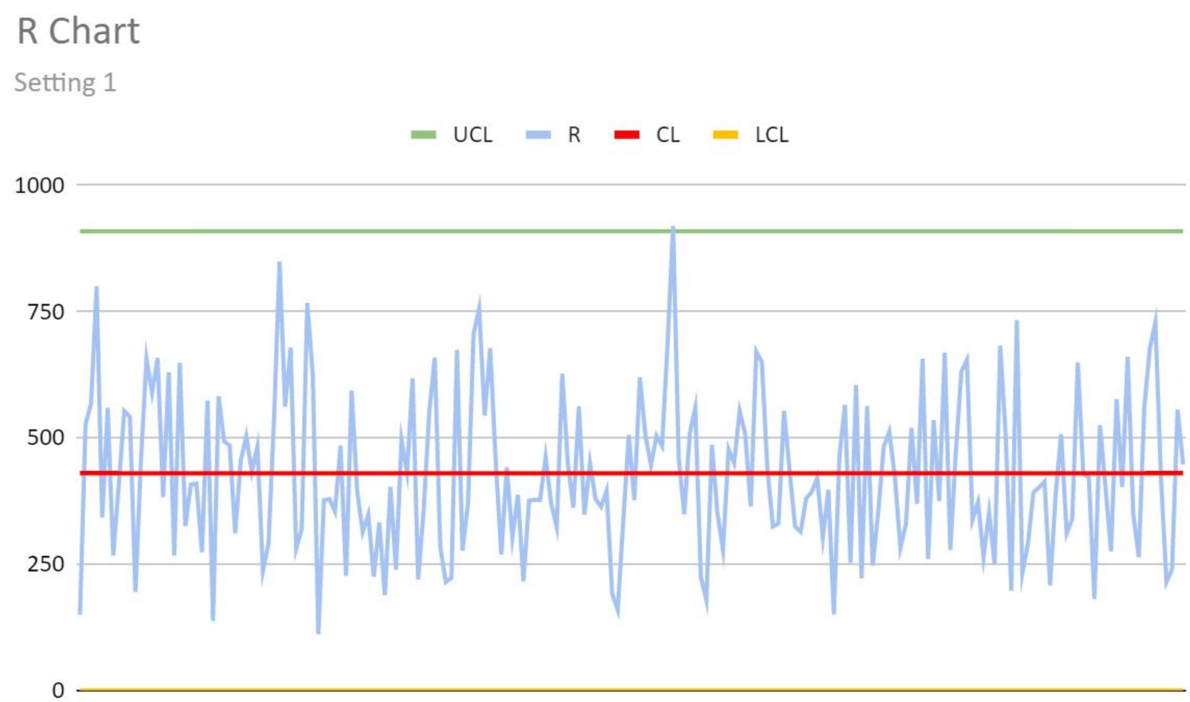

Figure 2.1: R Chart, Setting 1

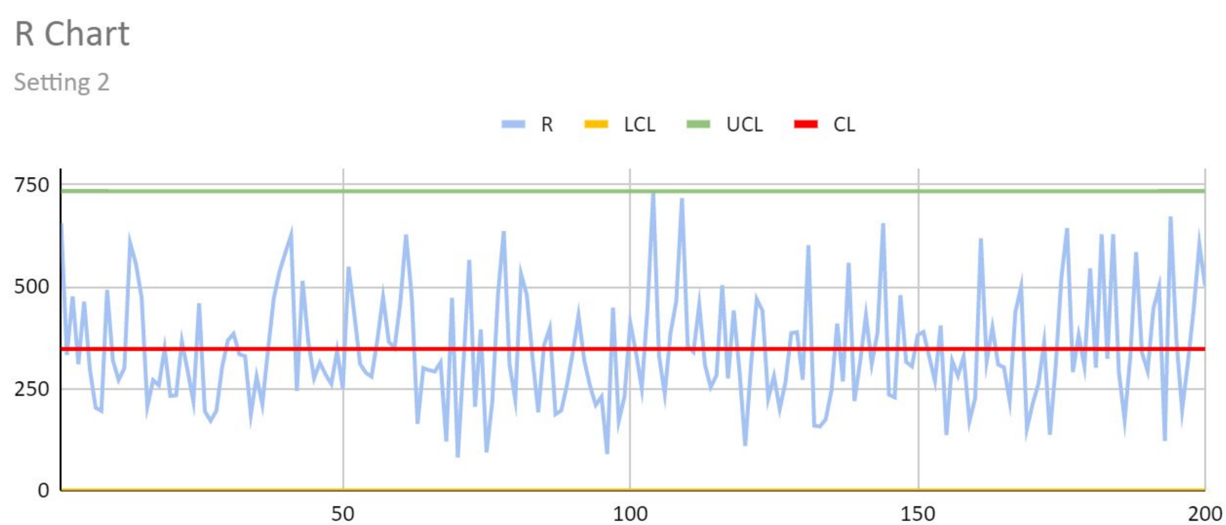

Figure 2.2: R Chart, Setting 2

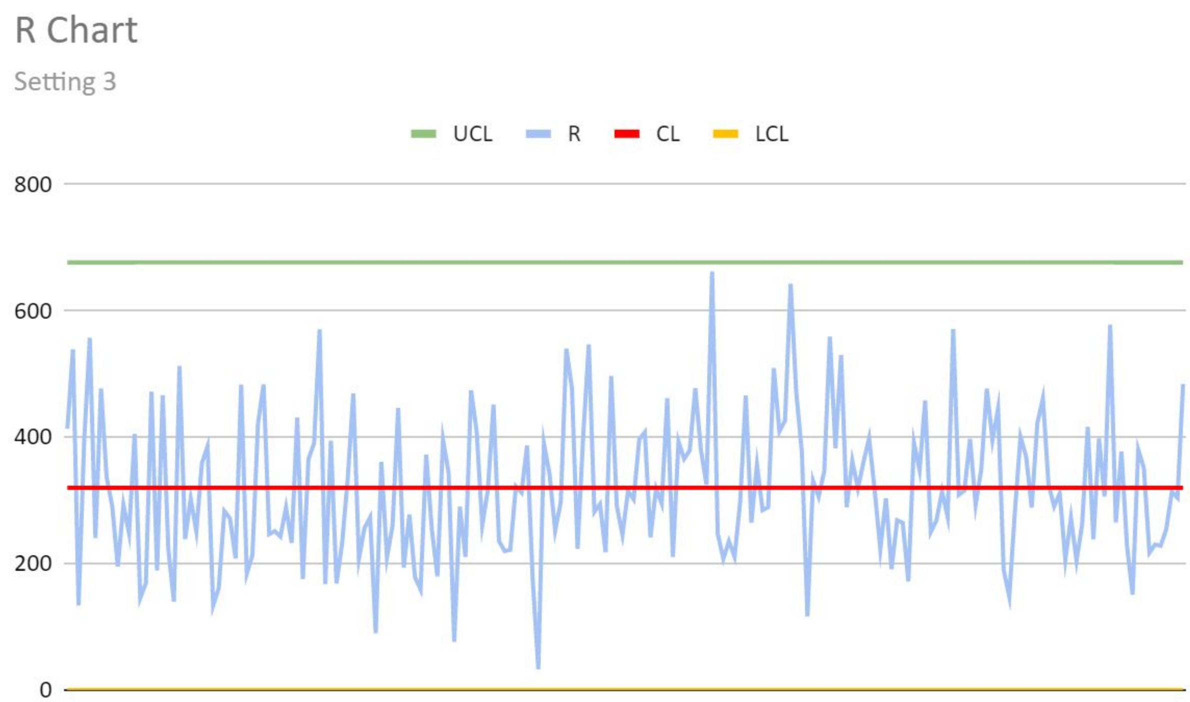

Figure 2.3: R Chart, Setting 3

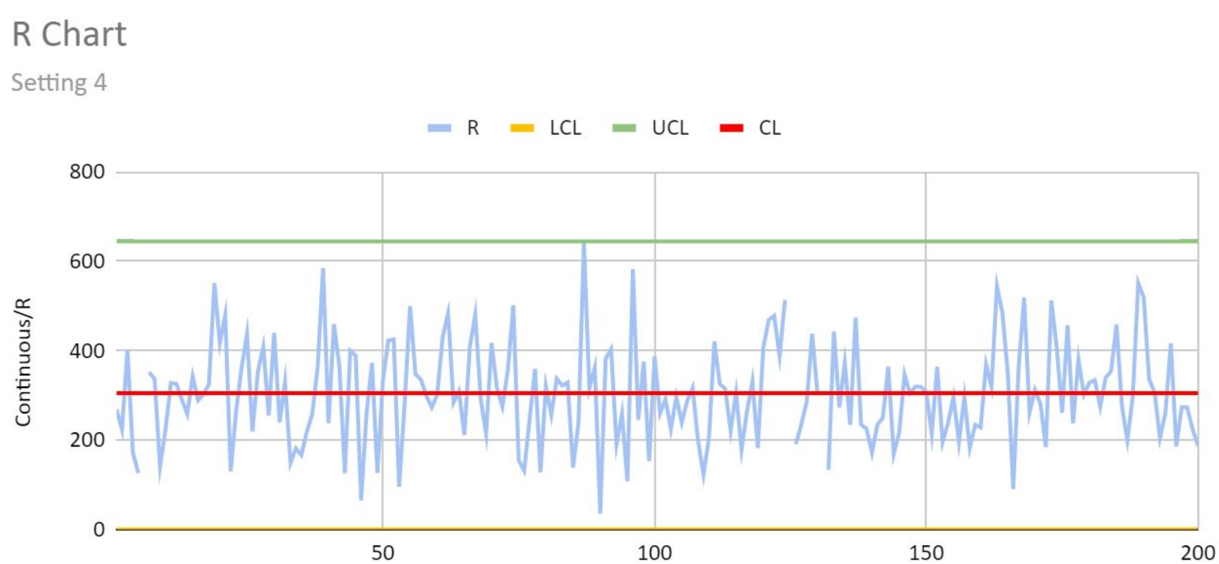

Figure 2.4: R Chart, Setting 4

<u>Time Series</u>

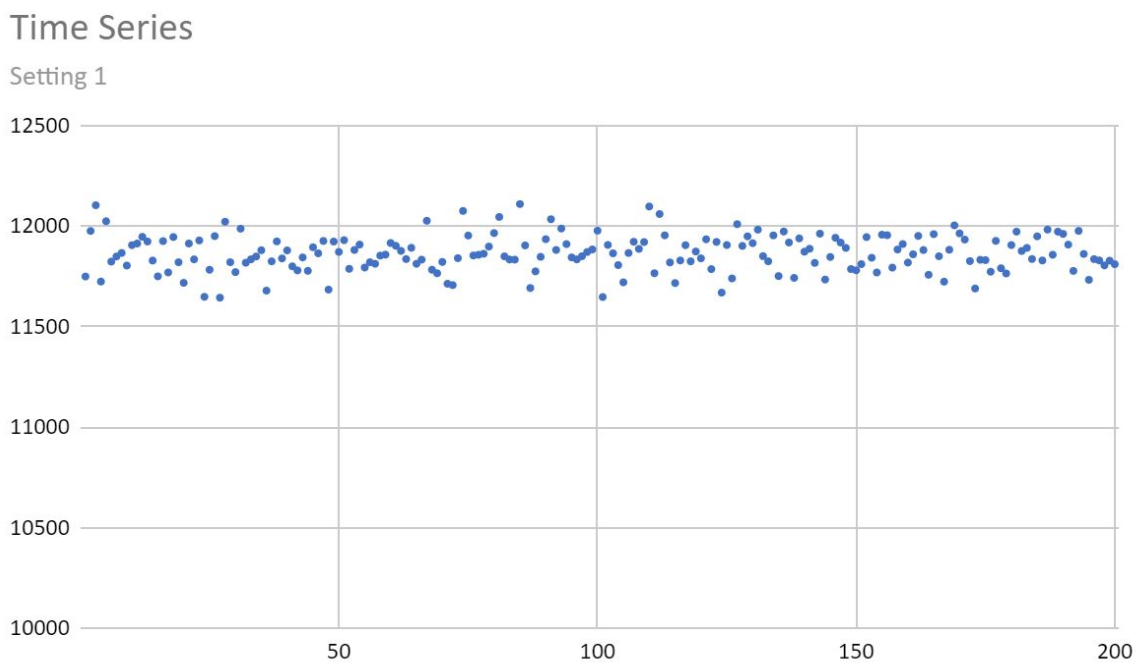

Figure 3.1: Time Series, Setting 1

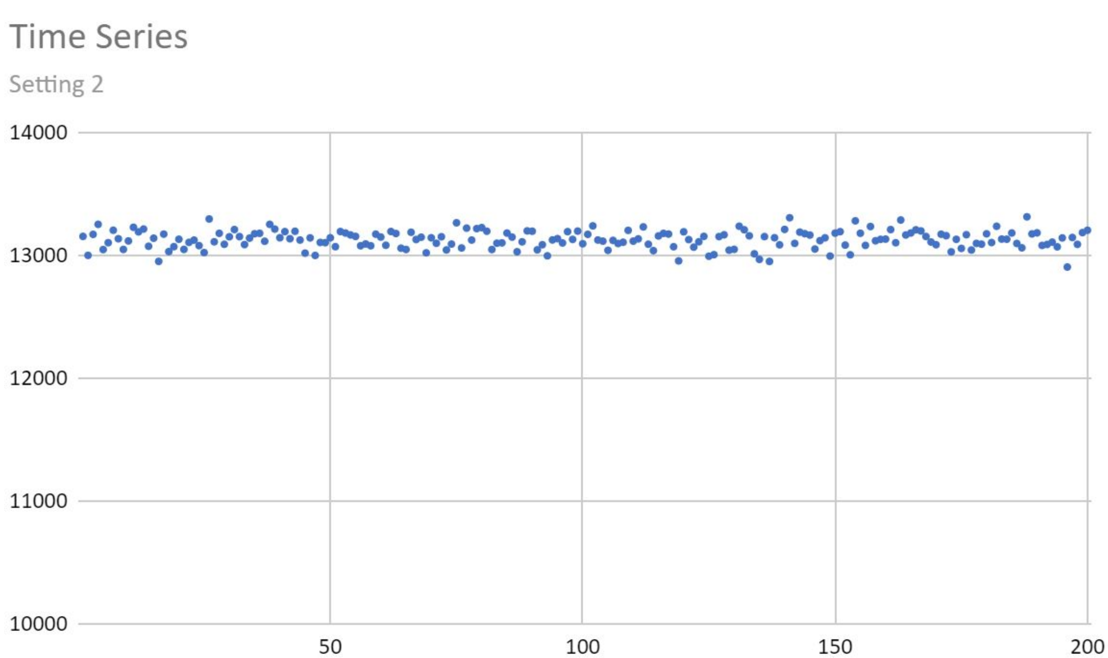

Figure 3.2: Time Series, Setting 2

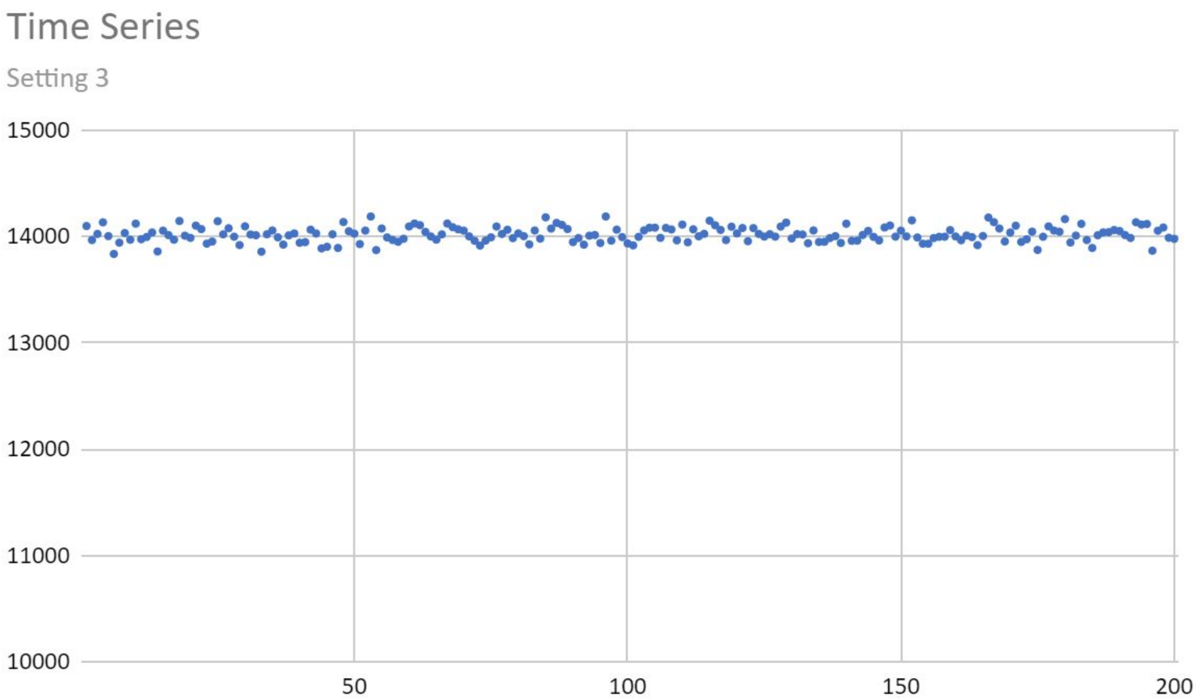

Figure 3.3: Time Series, Setting 3

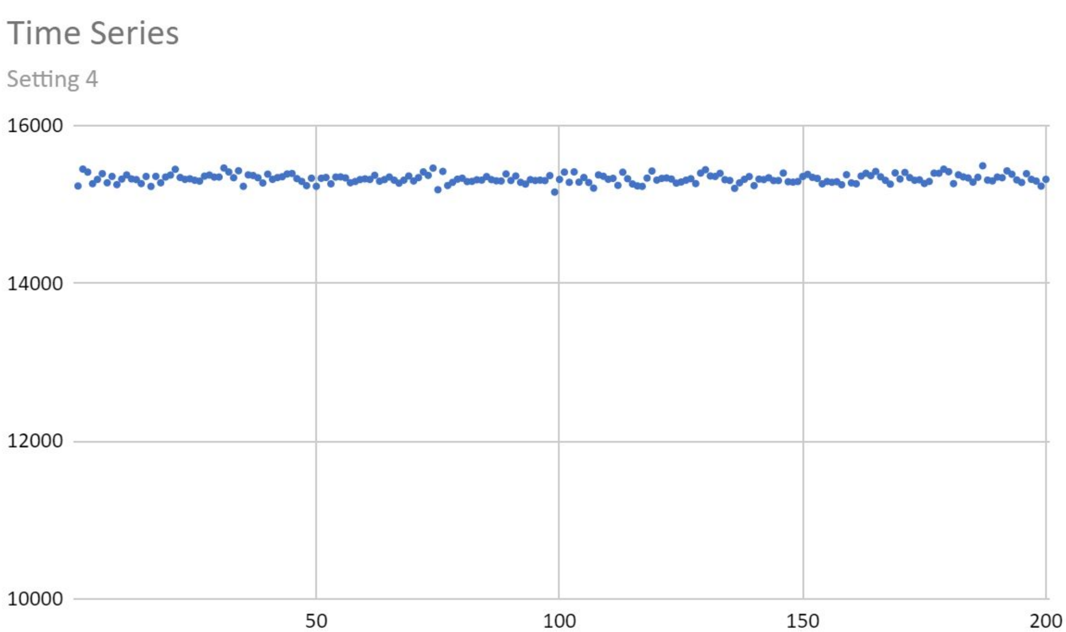

Figure 3.4: Time Series, Setting 4

Histogram

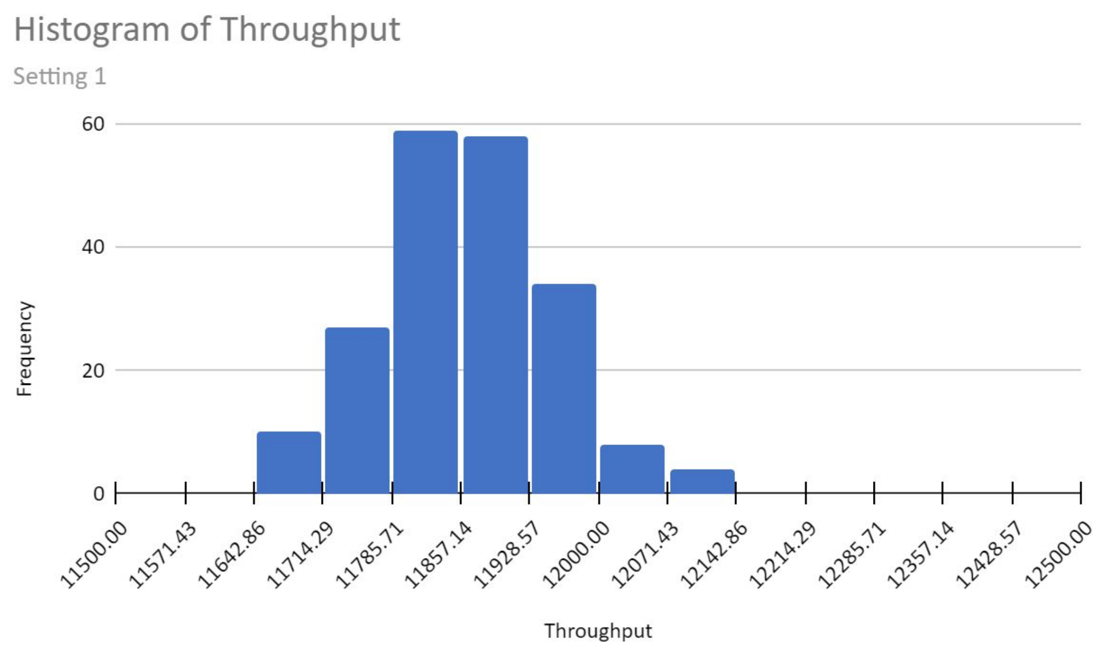

Figure 4.1: Histogram, Setting 1

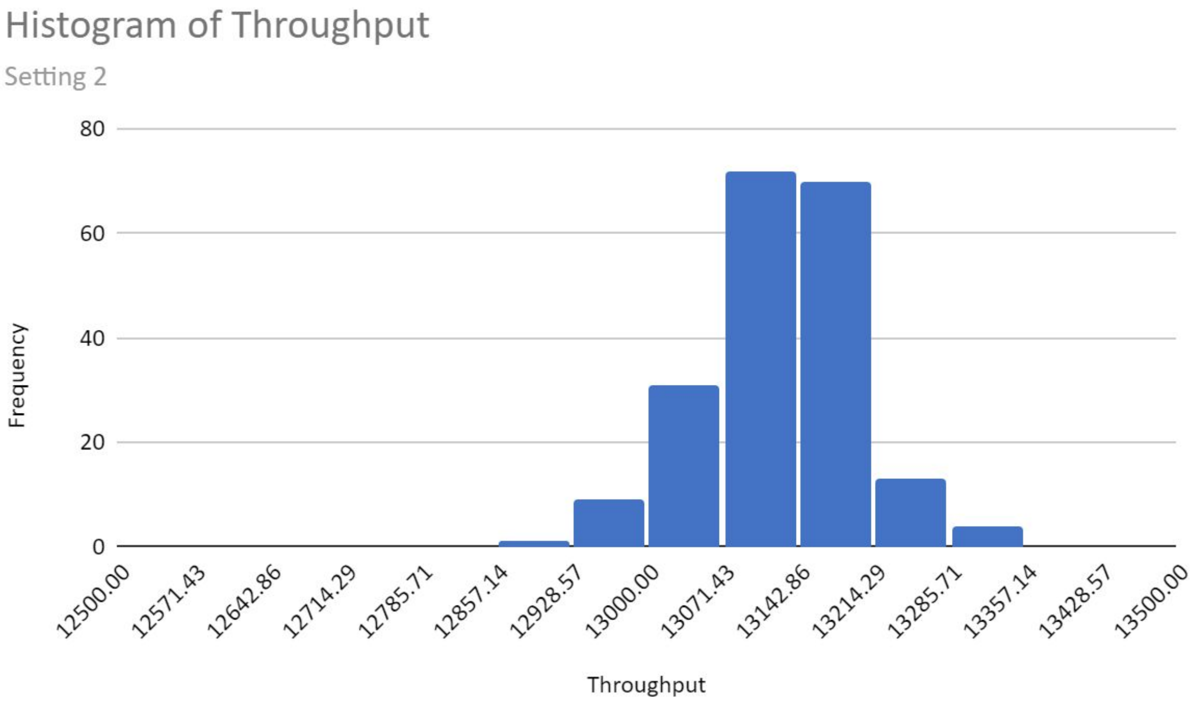

Figure 4.2: Histogram, Setting 2

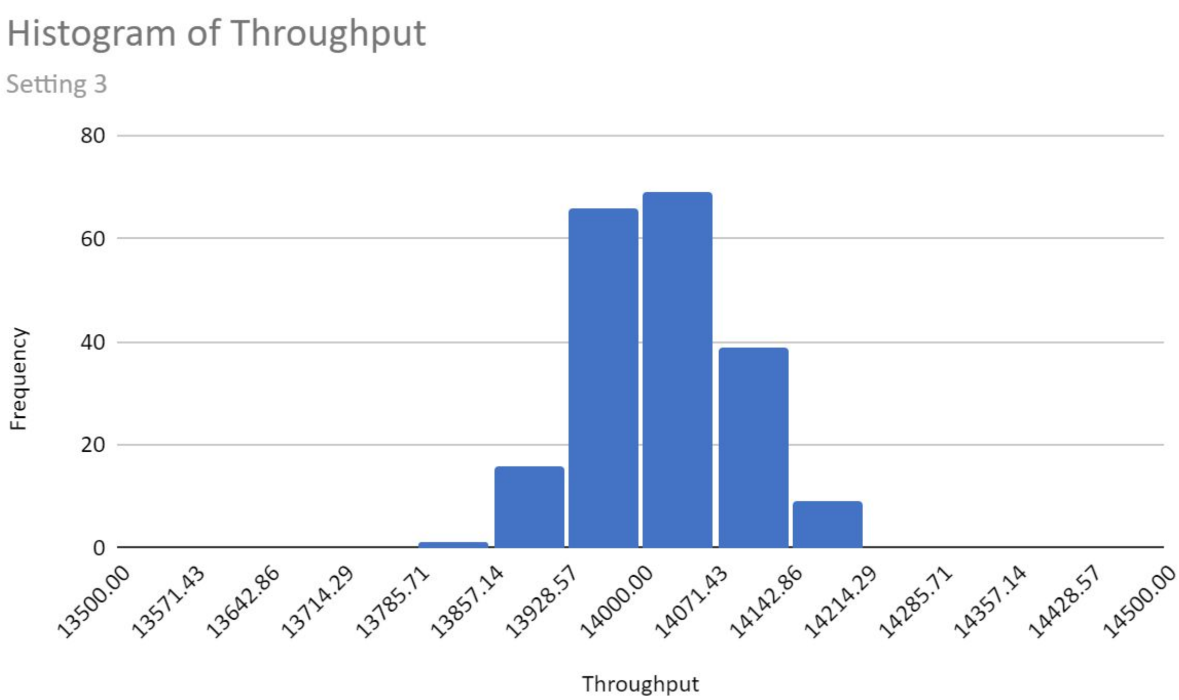

Figure 4.3: Histogram, Setting 3

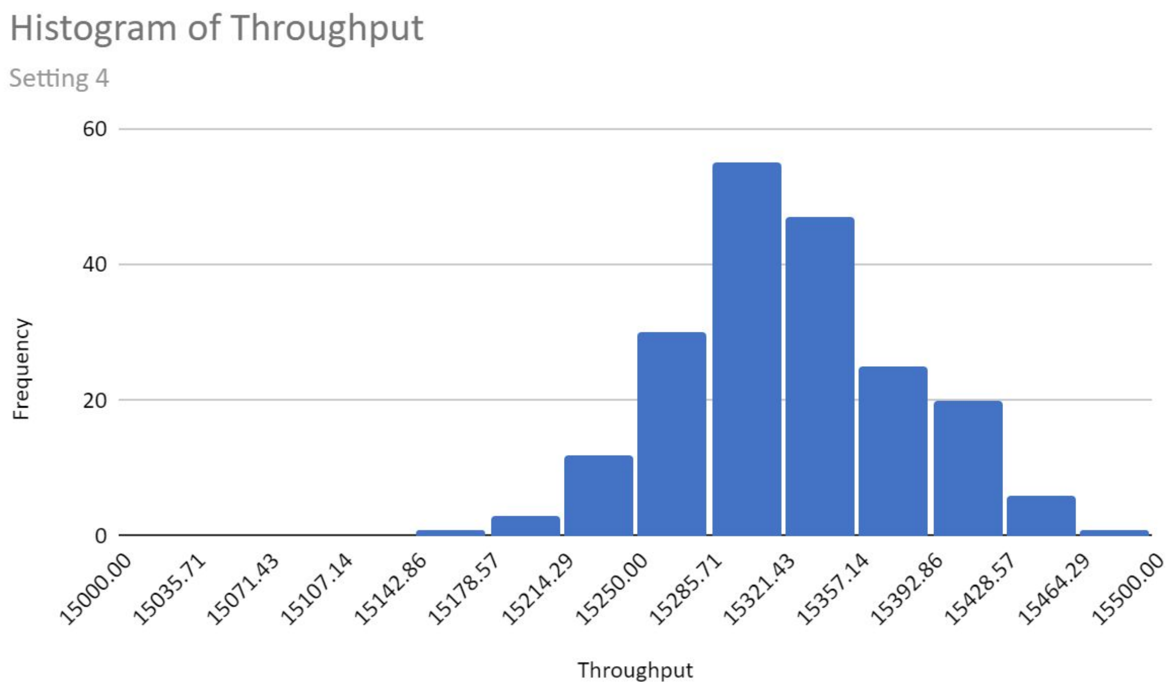

Figure 4.4: Histogram, Setting 4

NU

Histogram

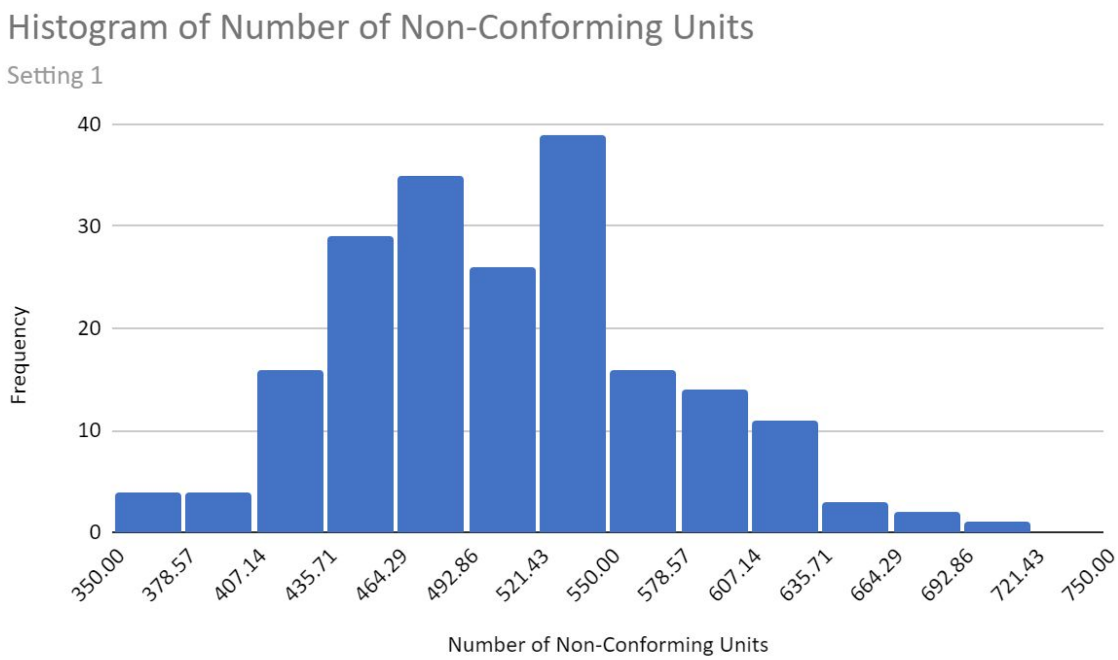

Figure 5.1: Histogram, Setting 1

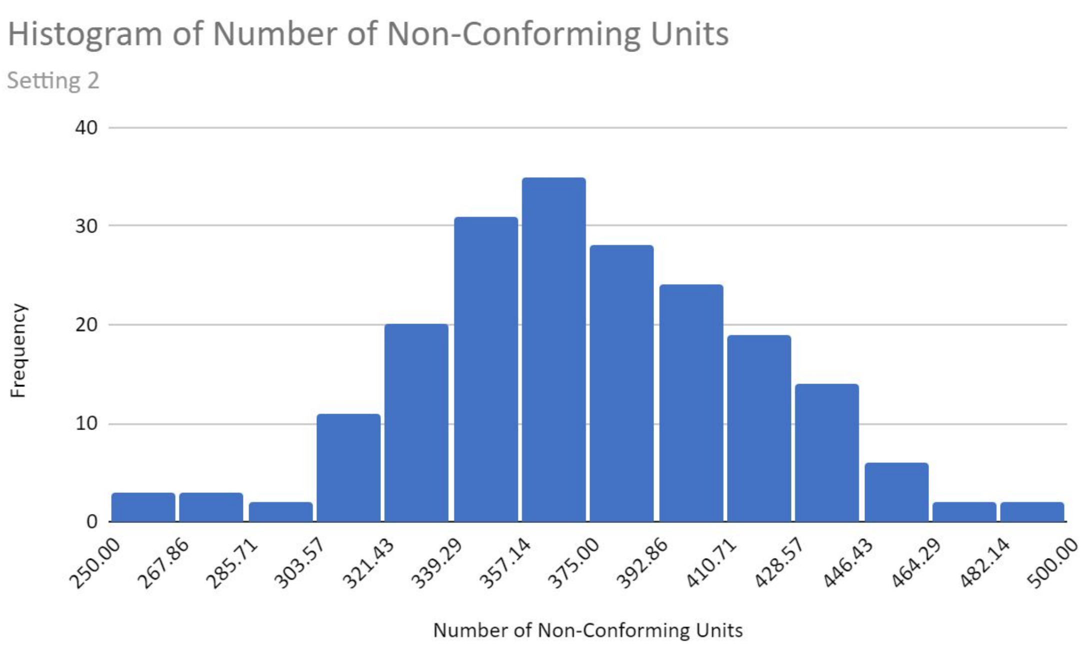

Figure 5.2: Histogram, Setting 2

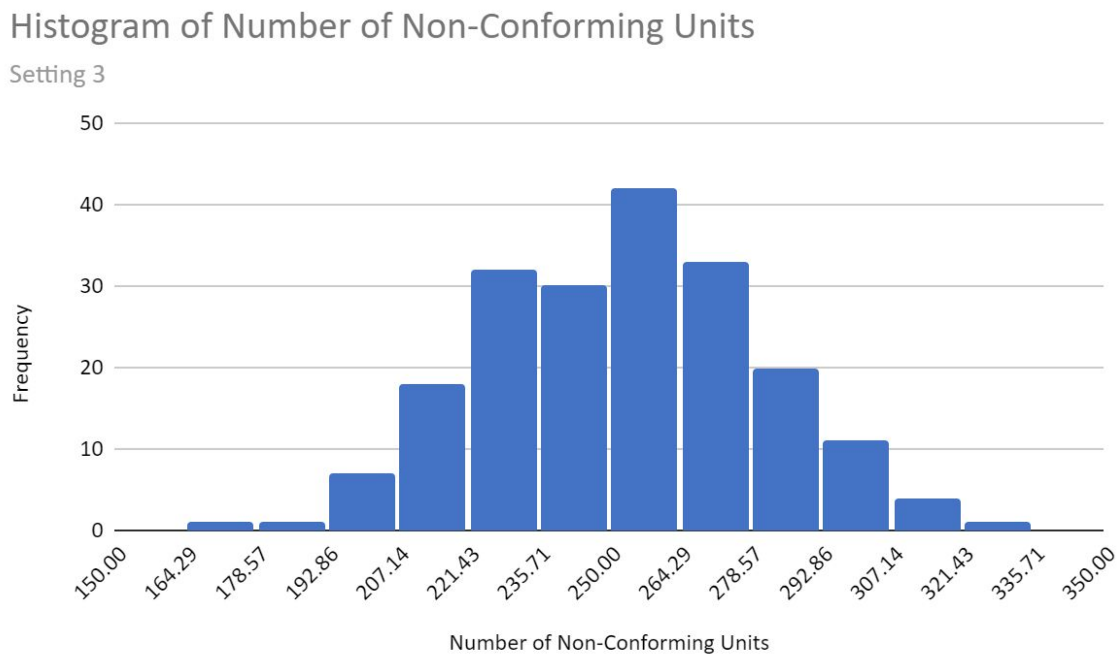

Figure 5.3: Histogram, Setting 3

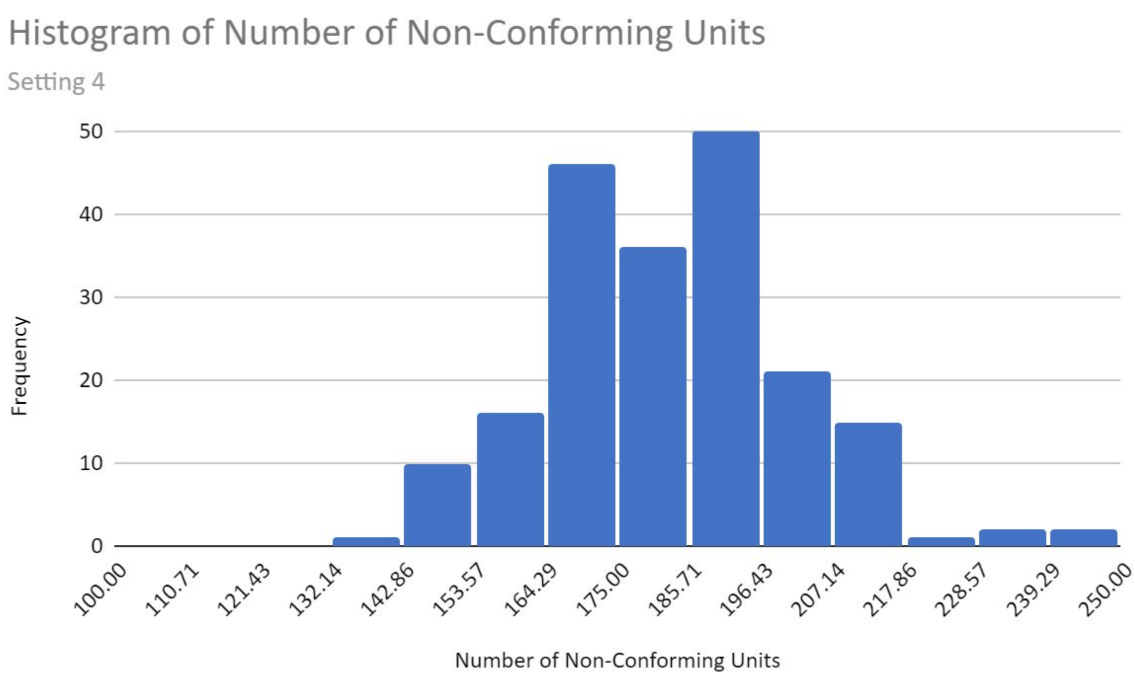

Figure 5.4: Histogram, Setting 4

<u>Box Plot</u>

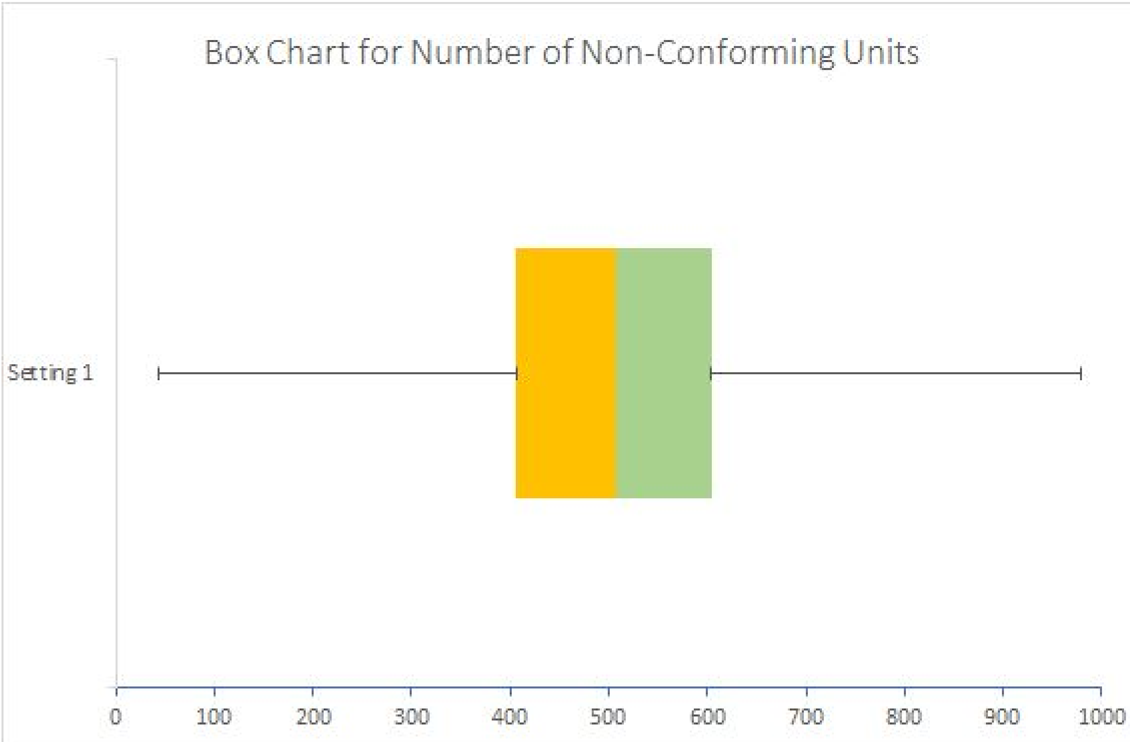

Figure 6.1: Box Plot, Setting 1

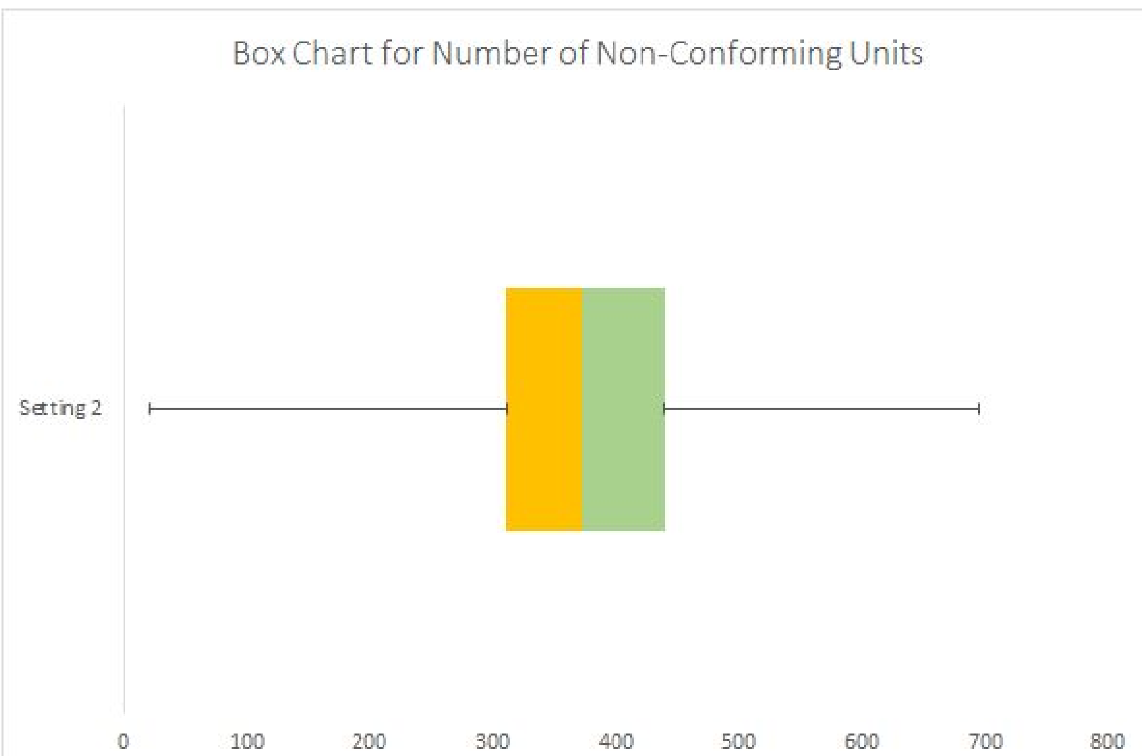

Figure 6.2: Box Plot, Setting 2

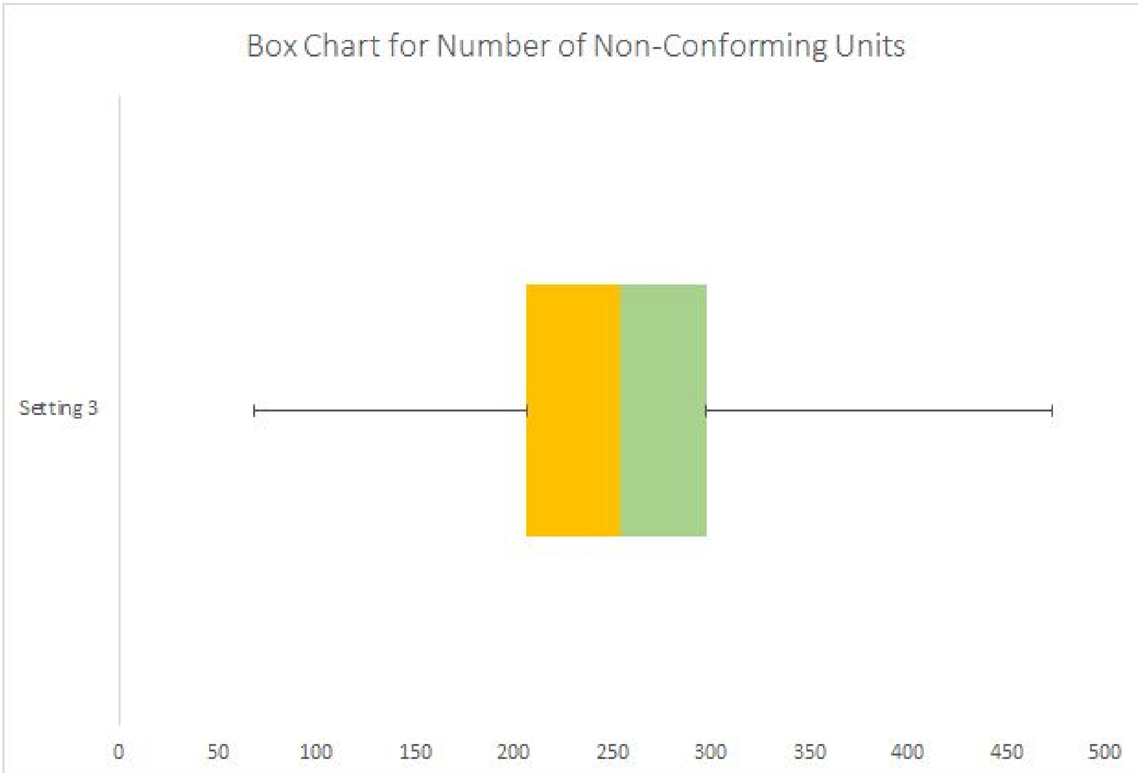

Figure 6.3: Box Plot, Setting 3

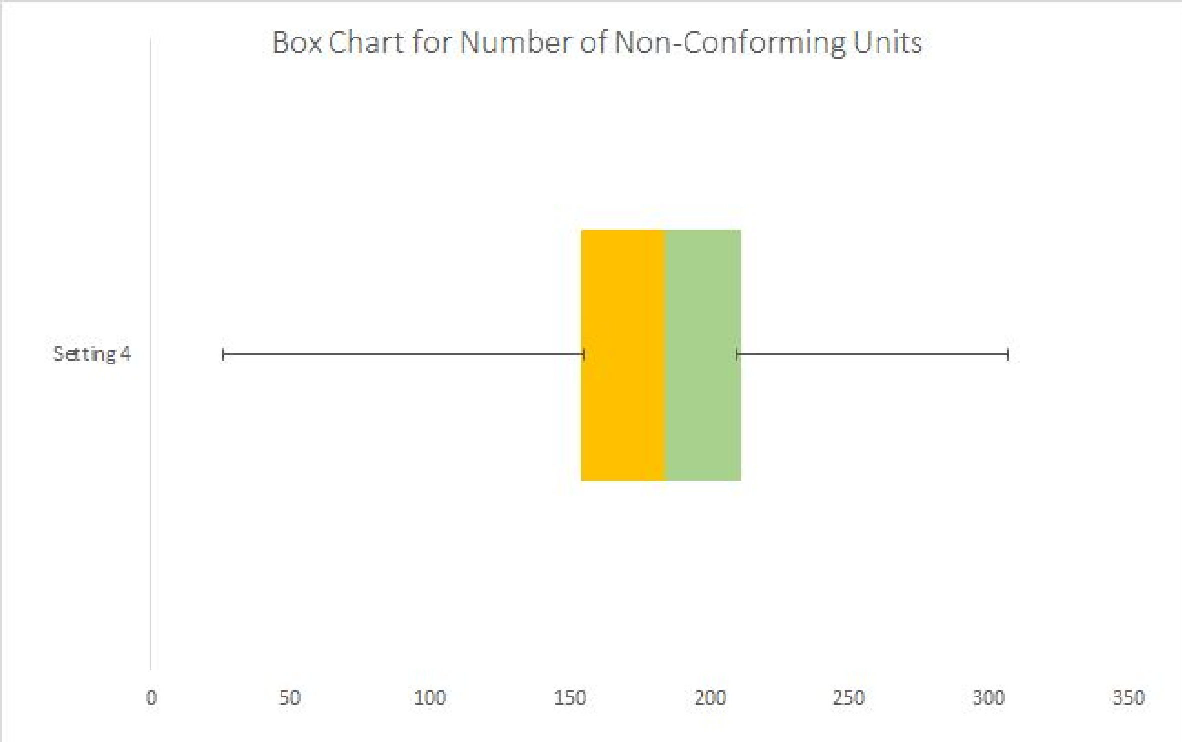

Figure 6.4: Box Plot, Setting 4

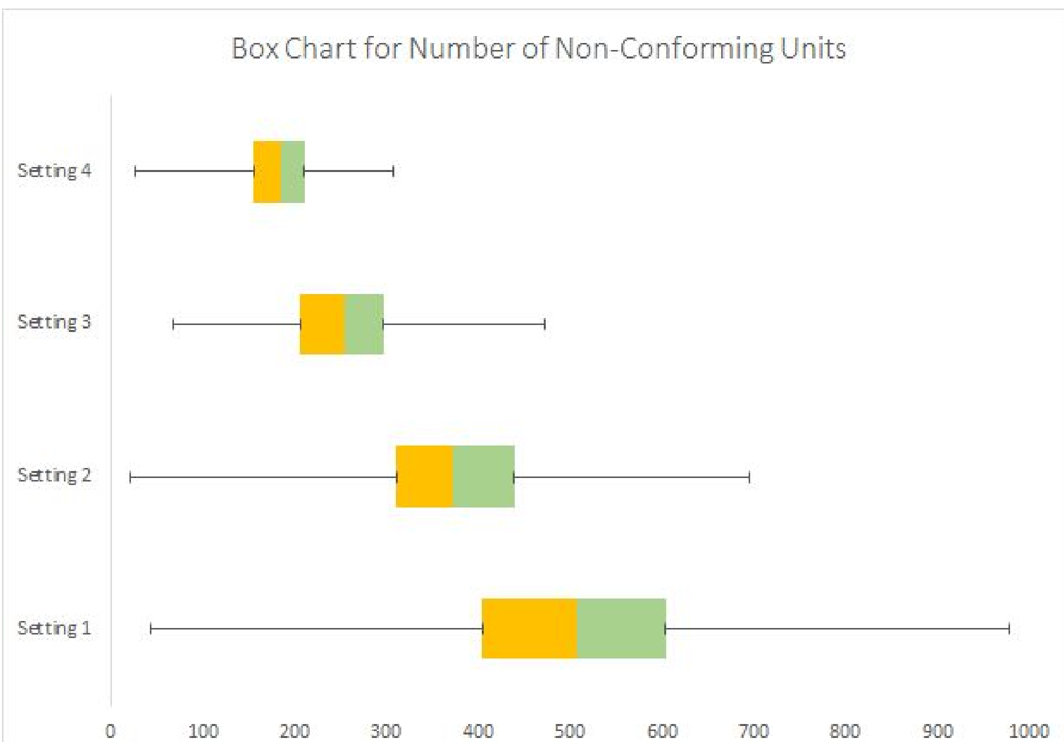

Figure 6.5: Box Plot, All Settings

P Chart

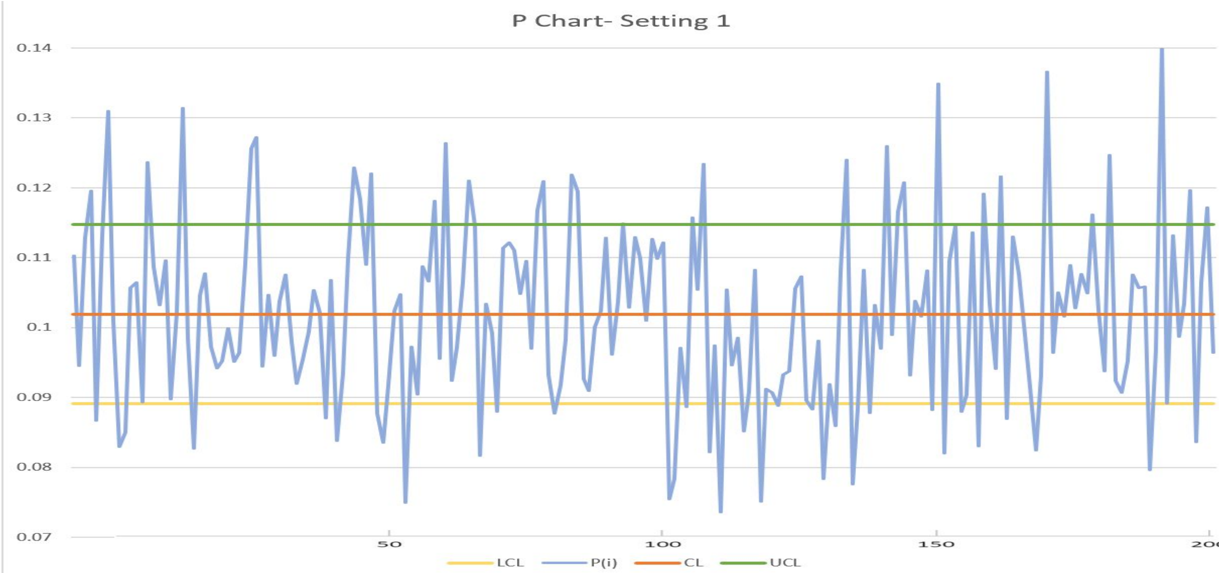

Figure 7.1: P Chart, Setting 1

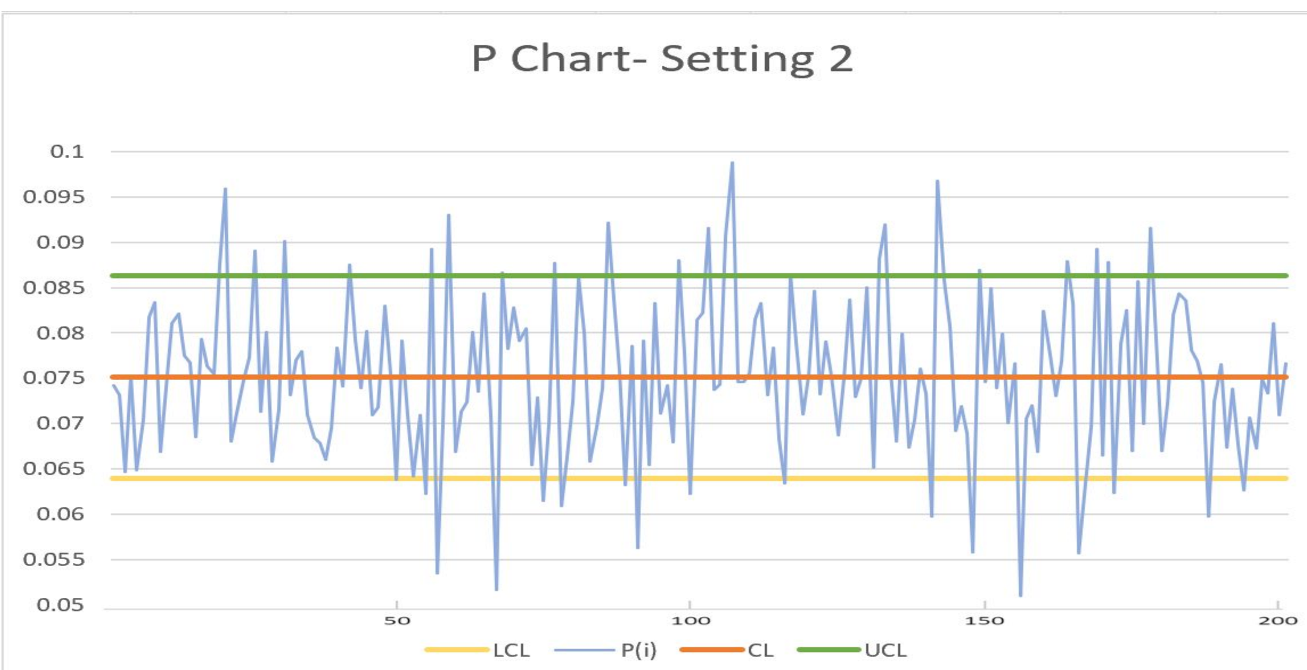

Figure 7.2: P Chart, Setting 2

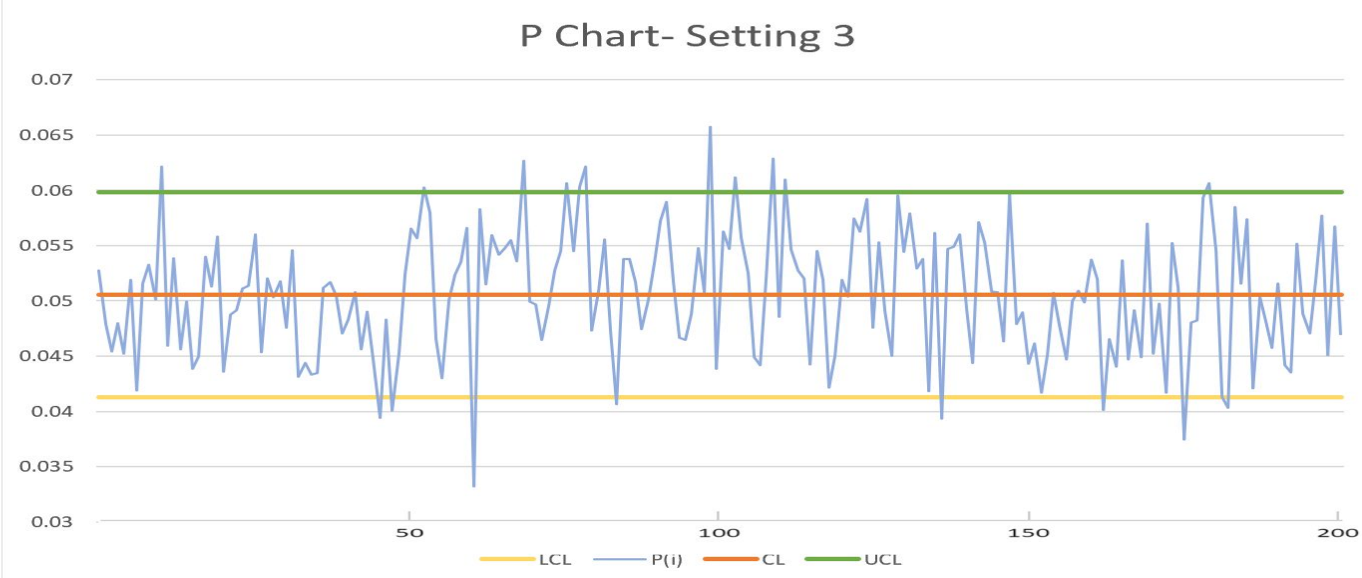

Figure 7.3: P Chart, Setting 3

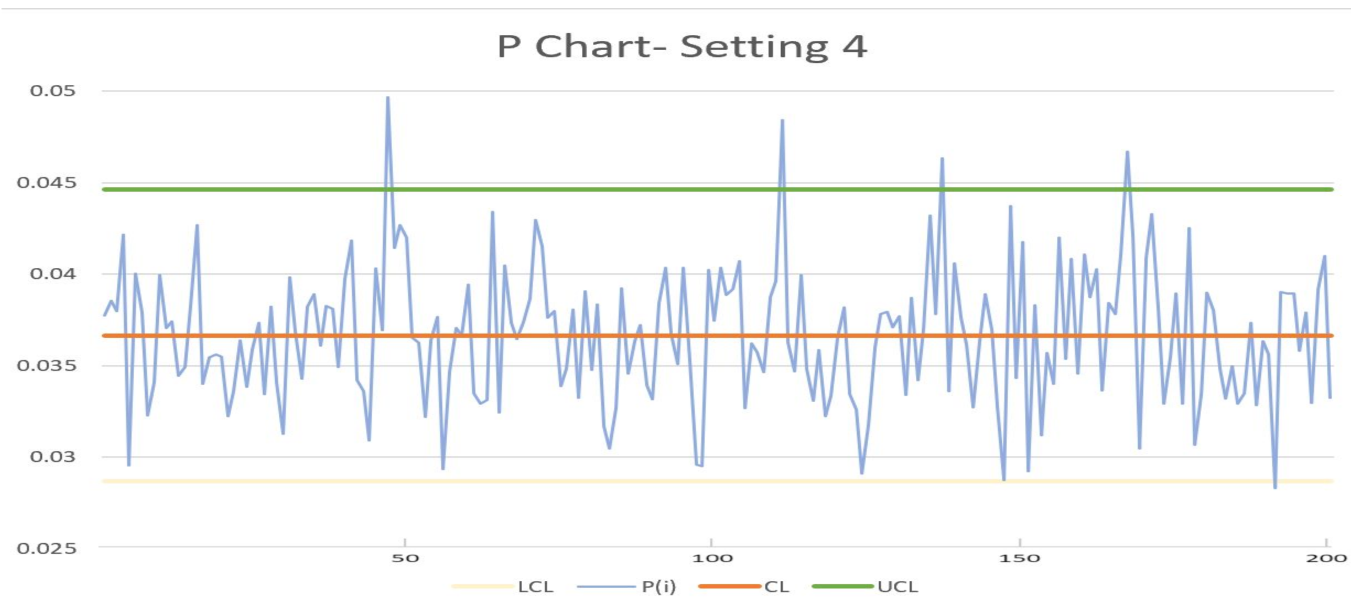

Figure 7.4: P Chart, Setting 4

Two Sample Hypothesis Testing for Throughput

Setting 1-2:

$H_0 : \mu_1 - \mu_2 = \Delta_0 = 0$

$H_1 : \mu_1 - \mu_2 \neq \Delta_0 = 0$

$S_1 = 89.948$

$S_2 = 71.930$

$\overline{x_1} = 11864.51103$

$\overline{x_2} = 13128.77586$

$$S_p^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$$

$$S_p = \sqrt{6632.3197} = 81.439$$

$$t_o = \frac{\bar{x}_1-\bar{x}_2-\Delta_0}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

$$t_o = \frac{11864.51103-13128.77586-0}{81.439\sqrt{\frac{1}{200}+\frac{1}{200}}} = -155.2407$$

$H_0 : \mu_1 - \mu_2 = \Delta_0 = 0$ will be rejected if $|(t_0)| > t_{\frac{\alpha}{2}, n_1+n_2-2}$

$$t_{\frac{\alpha}{2}, n_1+n_2-2} = t_{0.025, 398} = 1.96$$

Because $t_{0.025, 398} = 1.96$ and $|(t_0)| = 155.24$ the null hypothesis can be rejected at the significance level of 0.05 due to the sufficient evidence within the sample that favours the alternative hypothesis.

Setting 1-3:

$H_0 : \mu_1 - \mu_3 = \Delta_0 = 0$

$H_1 : \mu_1 - \mu_3 \neq \Delta_0 = 0$

$S_1 = 89.948$

$S_3 = 71.1588$

$\overline{x_1} = 11864.51103$

$\overline{x_3} = 14019.82$

$$S_p^2 = \frac{(n_1-1)S_1^2+(n_3-1)S_3^2}{n_1+n_3-2}$$

$$S_p = \sqrt{6577.1087} = 81.099$$

$$t_o = \frac{\bar{x}_1-\bar{x}_3-\Delta_0}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_3}}}$$

$$t_o = \frac{11864.51103-14019.82-0}{81.099\sqrt{\frac{1}{200}+\frac{1}{200}}} = -265.7627$$

$H_0 : \mu_1 - \mu_3 = \Delta_0 = 0$ will be rejected if $|(t_0)| > t_{\frac{\alpha}{2},\, n_1+n_2-2}$

$$t_{\frac{\alpha}{2},\, n_1+n_3-2} = t_{0.025,\, 398} = 1.96$$

Setting 1-4:

$H_0 : \mu_1 - \mu_4 = \Delta_0 = 0$

$H_1 : \mu_1 - \mu_4 \neq \Delta_0 = 0$

$S_1 = 89.948$

$S_4 = 56.263$

$\overline{x_1} = 11864.51103$

$\overline{x_4} = 15325.884$

$$S_p^2 = \frac{(n_1-1)S_1^2+(n_4-1)S_4^2}{n_1+n_4-2}$$

$$S_p = \sqrt{5628.0895} = 75.020$$

$$t_o = \frac{\bar{x}_1-\bar{x}_4-\Delta_0}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_4}}}$$

$$t_o = \frac{11864.51103-15325.884-0}{75.020\sqrt{\frac{1}{200}+\frac{1}{200}}} = -461.389$$

$H_0 : \mu_1 - \mu_4 = \Delta_0 = 0$ will be rejected if $|(t_0)| > t_{\frac{\alpha}{2}, n_1+n_2-2}$

$$t_{\frac{\alpha}{2}, n_1+n_4-2} = t_{0.025, 398} = 1.96$$

Setting 2-3:

$H_0 : \mu_2 - \mu_3 = \Delta_0 = 0$

$H_1 : \mu_2 - \mu_3 \neq \Delta_0 = 0$

$S_2 = 71.930$

$S_3 = 71.1588$

$\overline{x_2} = 13128.77586$

$\overline{x_3} = 14019.82$

$$S_p^2 = \frac{(n_2-1)S_2^2+(n_3-1)S_3^2}{n_2+n_3-2}$$

$$S_p = \sqrt{5118.749} = 71.545$$

$$t_o = \frac{\bar{x}_2 - \bar{x}_3 - \Delta_0}{S_p\sqrt{\frac{1}{n_2}+\frac{1}{n_3}}}$$

$$t_o = \frac{13128.77586 - 14019.82 - 0}{71.545\sqrt{\frac{1}{200}+\frac{1}{200}}} = -124.542$$

$H_0 : \mu_2 - \mu_3 = \Delta_0 = 0$will be rejected if $|(t_0)| > t_{\frac{\alpha}{2},\, n_1+n_2-2}$

$t_{\frac{\alpha}{2},\, n_2+n_3-2} = t_{0.025,\, 398} = 1.96$

Setting 2-4:

$H_0 : \mu_2 - \mu_4 = \Delta_0 = 0$

$H_1 : \mu_2 - \mu_4 \neq \Delta_0 = 0$

$S_2 = 71.930$

$S_4 = 56.263$

$\overline{x_2} = 13128.77586$

$\overline{x_4} = 15325.884$

$$S_p^2 = \frac{(n_2-1)S_2^2+(n_4-1)S_4^2}{n_2+n_4-2}$$

$$S_p = \sqrt{4169.7250} = 64.5734$$

$$t_o = \frac{\bar{x}_2 - \bar{x}_4 - \Delta_0}{S_p\sqrt{\frac{1}{n_2}+\frac{1}{n_4}}}$$

$$t_o = \frac{13128.77586 - 15325.884 - 0}{64.5734\sqrt{\frac{1}{200}+\frac{1}{200}}} = -340.249$$

$H_0 : \mu_2 - \mu_4 = \Delta_0 = 0$ will be rejected if $|(t_0)| > t_{\frac{\alpha}{2},\, n_1+n_2-2}$

$$t_{\frac{\alpha}{2},\, n_2+n_4-2} = t_{0.025,\, 398} = 1.96$$

Setting 3-4:

$H_0 : \mu_3 - \mu_4 = \Delta_0 = 0$

$H_1 : \mu_3 - \mu_4 \neq \Delta_0 = 0$

$S_3 = 71.1588$

$S_4 = 56.263$

$\overline{x_3} = 14019.82$

$\overline{x_4} = 15325.884$

$$S_p^2 = \frac{(n_3-1)S_3^2+(n_4-1)S_4^2}{n_3+n_4-2}$$

$$S_p = \sqrt{4114.549} = 64.144$$

$$t_o = \frac{\bar{x}_3-\bar{x}_4-\Delta_0}{S_p\sqrt{\frac{1}{n_3}+\frac{1}{n_4}}}$$

$$t_o = \frac{14019.82-15325.884-0}{64.144\sqrt{\frac{1}{200}+\frac{1}{200}}} = -203.708$$

$H_0 : \mu_3 - \mu_4 = \Delta_0 = 0$ will be rejected if $|(t_0)| > t_{\frac{\alpha}{2},\, n_1+n_2-2}$

$$t_{\frac{\alpha}{2},\, n_3+n_4-2} = t_{0.025,\, 398} = 1.96$$

<u>Z- test for number of non conforming units</u>

$H_0 :\ p_1 = p_2$

$H_1 :\ p_1 > p_2$

Sample 1-2:

$p_1 = 0.101884$

$p_2 = 0.075124$

$n_1 = 5000$

$n_2 = 5000$

p_hat= $\frac{n1p1-n2p2}{n1+n2} = 0.01338$

$Z_o = \frac{p1-p2}{/phat(1-phat)(\frac{1}{n1}+\frac{1}{n2})} = 11.64$

$Z_o = 69.20 \; > \; Z_{0.05,} = 1.645$

Reject H_o

Sample 1-3:

$p_1 = 0.101884$

$p_2 = 0.050568$

n1=5000

n2=5000

p_hat= $\frac{n1p1-n2p2}{n1+n2} = 0.025658$

$Z_o = \frac{p1-p2}{/phat(1-phat)(\frac{1}{n1}+\frac{1}{n2})} = 16.227$

$Z_o = 16.227 \; > \; Z_{0.05,} = 1.645$

Reject H_o

Sample 1-4:

$p_1 = 0.101884$

$p_2 = 0.036651$

$n_1 = 5000$

$n_2 = 5000$

p_hat= $\frac{n1p1-n2p2}{n1+n2} = 0.03262$

$Z_o = \frac{p1-p2}{/phat(1-phat)(\frac{1}{n1}+\frac{1}{n2})} = 18.36$

$Z_o = 18.36 \; > \; Z_{0.05,} = 1.645$

Reject H_o

Sample 2-3:

p_1 =0.075124

p_2 =0.050568

n_1 =5000

n_2 =5000

p_hat= $\frac{n1p1-n2p2}{n1+n2}$ = 0.012278

$Z_o = \frac{p1-p2}{\sqrt{phat(1-phat)(\frac{1}{n1}+\frac{1}{n2})}} = 11.10$

$Z_o = 11.10 > Z_{0.05,} = 1.645$

Reject H_o

Sample 2-4:

p_1 =0.075124

p_2 = 0.036651

n_1 =5000

n_2 =5000

p_hat= $\frac{n1p1-n2p2}{n1+n2}$ = 0.01923

$Z_o = \frac{p1-p2}{\sqrt{phat(1-phat)(\frac{1}{n1}+\frac{1}{n2})}} = 14.01$

$Z_o = 14.01 > Z_{0.05,} = 1.645$

Reject H_o

Sample 3-4:

p_1 =0.050568

p_2 = 0.036651

n_1 =5000

n_2 =5000

p_hat= $\frac{n1p1-n2p2}{n1+n2}$ = 0.006959

$$Z_o = \frac{p1-p2}{\sqrt{phat(1-phat)(\frac{1}{n1}+\frac{1}{n2})}} = 8.37$$

$Z_o = 8.37 > Z_{0.05,} = 1.645$

Reject H_o

ANOVA for NU

SUMMARY

Groups	Count	Sum	Average	Variance
Row 1	5	2369	473.8	8350.7
Row 2	5	1924	384.8	10351.7
Row 3	5	1268	253.6	4346.8
Row 4	5	1083	216.6	2234.3

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	211793.2	3	70597.73	11.16898	0.000335	3.238872
Within Groups	101134	16	6320.875			
Total	312927.2	19				

ANOVA for TP

SUMMARY

Groups	Count	Sum	Average	Variance
Row 1	5	58999.5	11799.9	51177.35
Row 2	5	65409.77	13081.95	5598.653
Row 3	5	69905.08	13981.02	16165.33
Row 4	5	76678.27	15335.65	22688.18

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	33281275	3	11093758	464.0307	9.23E-16	3.238872
Within Groups	382518.1	16	23907.38			
Total	33663793	19				