ESCA PROJECT

27257983

Liam Stander

Contents

Introduction 2
Part 1 3
Customer data 3
Sales 3
Product head office 4
Products data 4
Suggestions 5
Part 3 6
Overall SPC Summary (All Products) 6
Product: CLO 6
Process Capability Summary 6
X̄ (Mean) Control Chart – Initial 30 Samples 7
S (Standard Deviation) Control Chart – Initial 30 Samples 8
X̄ (Mean) Control Chart – Monitoring Samples (31+) 8
S (Standard Deviation) Control Chart – Monitoring Samples (31+) 9
Insights and Interpretation 9
Product: KEY 10
Process Capability Summary 10
X̄ (Mean) Control Chart – Initial 30 Samples 10
S (Standard Deviation) Control Chart – Initial 30 Samples 11
X̄ (Mean) Control Chart – Monitoring Samples (31+) 11
S (Standard Deviation) Control Chart – Monitoring Samples (31+) 12
Insights and Interpretation 12
Product: LAP 13
Process Capability Summary 13
X̄ (Mean) Control Chart – Initial 30 Samples 13
S (Standard Deviation) Control Chart – Initial 30 Samples 14
X̄ (Mean) Control Chart – Monitoring Samples (31+) 14
S (Standard Deviation) Control Chart – Monitoring Samples (31+) 15
Insights and Interpretation 15
Product: MON 16
Process Capability Summary 16
X̄ (Mean) Control Chart – Initial 30 Samples 16
S (Standard Deviation) Control Chart – Initial 30 Samples 17
X̄ (Mean) Control Chart – Monitoring Samples (31+) 17
S (Standard Deviation) Control Chart – Monitoring Samples (31+) 18

Insights and Interpretation 18
Product: MOU 19
Process Capability Summary 19
X̄ (Mean) Control Chart – Initial 30 Samples 19
S (Standard Deviation) Control Chart – Initial 30 Samples 20
X̄ (Mean) Control Chart – Monitoring Samples (31+) 20
S (Standard Deviation) Control Chart – Monitoring Samples (31+) 21
Insights and Interpretation 21
Product: SOF 22
Process Capability Summary 22
X̄ (Mean) Control Chart – Initial 30 Samples 22
S (Standard Deviation) Control Chart – Initial 30 Samples 23
X̄ (Mean) Control Chart – Monitoring Samples (31+) 23
S (Standard Deviation) Control Chart – Monitoring Samples (31+) 24
Insights and Interpretation 24
Overall Capability Summary 25
Part4 26
4.1 Type 1 error 26
4.2 Type 2 error 27
Part 4.3 updated data analysis 28
Part 5 Baristas and time to serve 29
Summary by Number of Baristas shop 1 29
Optimal Staffing Level 30
Insights and Interpretation 30
Summary by Number of Baristas for shop 2 31
Optimal Staffing Level 32
Insights and Interpretation 32
Part 6 ANOVA 33
Part 7.1 34
Part 7.2 35
Conclusion 36

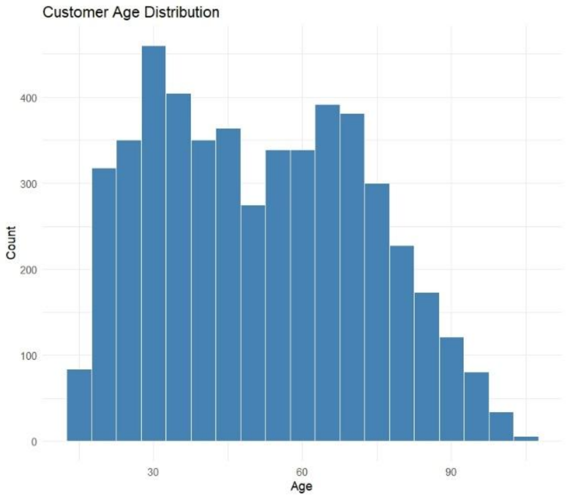
Introduction

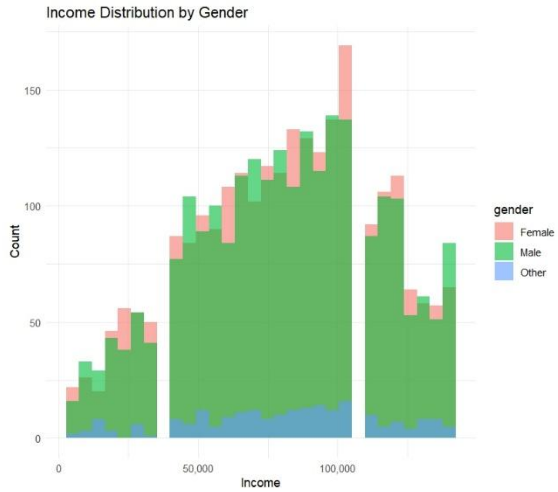
This ESCA project is about business data and to give them back the feedback needed to improve efficiency and productivity, grow financially and make the best decisions for the business to flourish. I will be giving back that feedback on given data sets where I have done the necessary tests and procedures to enlighten them.

Part 1

Customer data

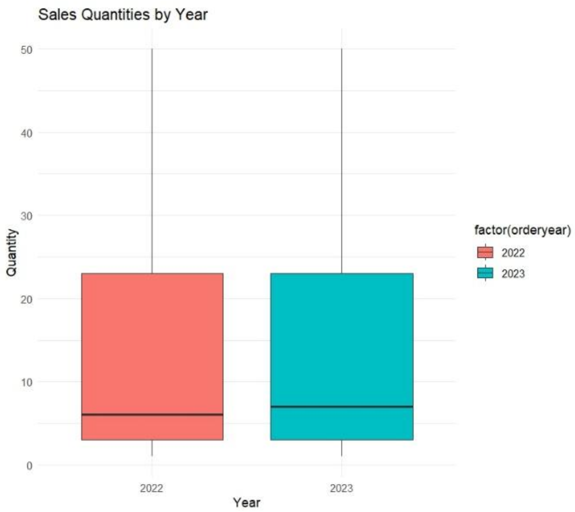
The following 2 graphs show us the distribution of customers with the age majority between 20 and 40 thus younger target audience. And there are slightly more female customers then male. Income varies thus potential for premium segmentation.

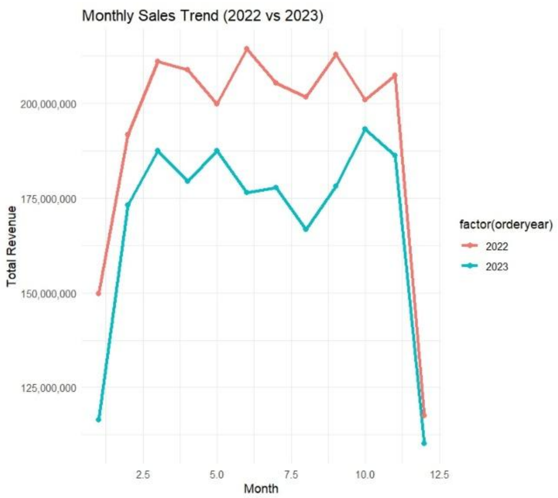




Sales

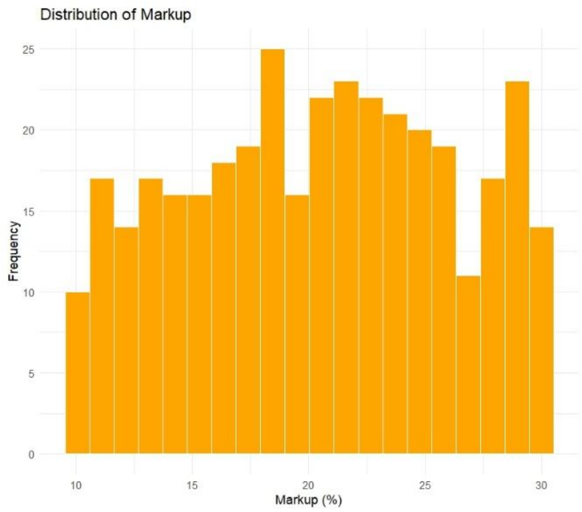
From the sales data we can see that there is no major difference in sales over the 2-year period with a small increase in mean sales and thus positive growth. The monthly sales, however, has dropped quite a bit, which is not a good sign for the company, but this could be because of a decrease in average selling price since the quantity is consistent. Monthly peaks can be seen thus suggest seasonal demand which can be leveraged by forecasting and inventory planning.

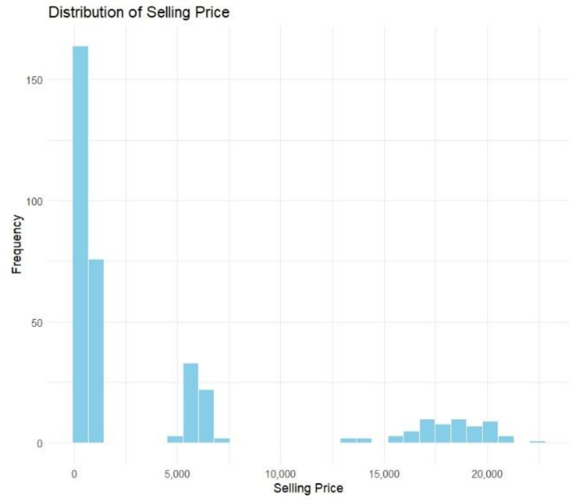


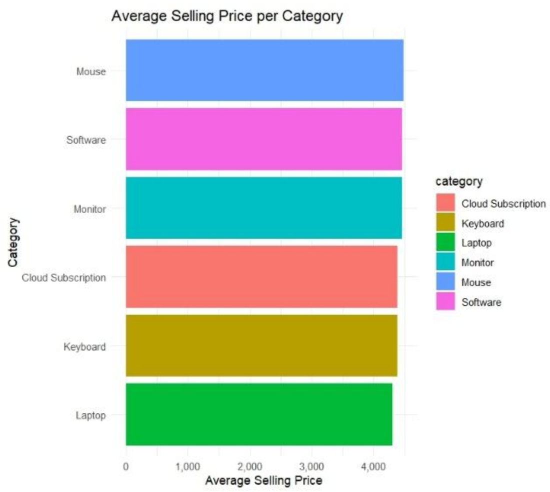


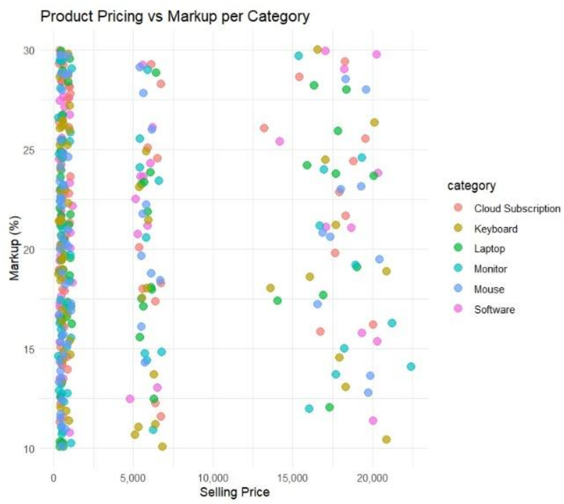
Product head office

The markup is widely distributed as well as the selling price. The pricing vs markup shows 3 major pricing ranges and the majority selling prices are in the lower range. The average selling price for each category is also very consistent.



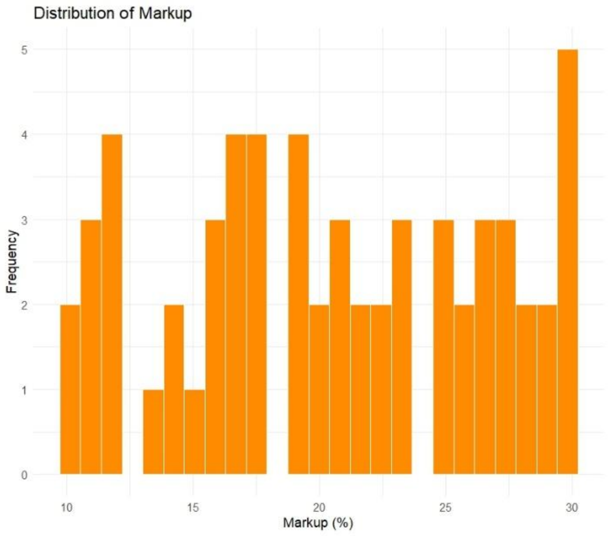


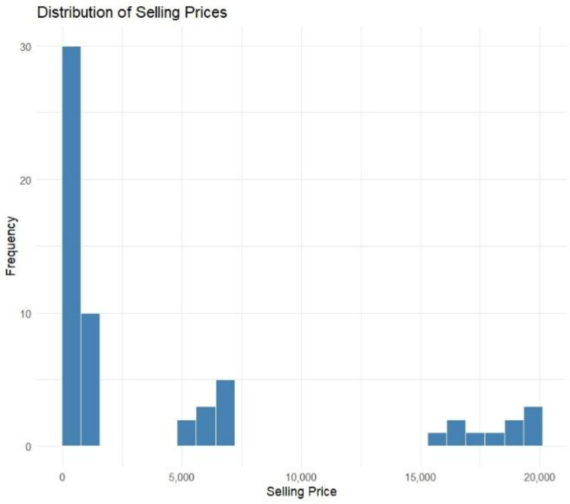


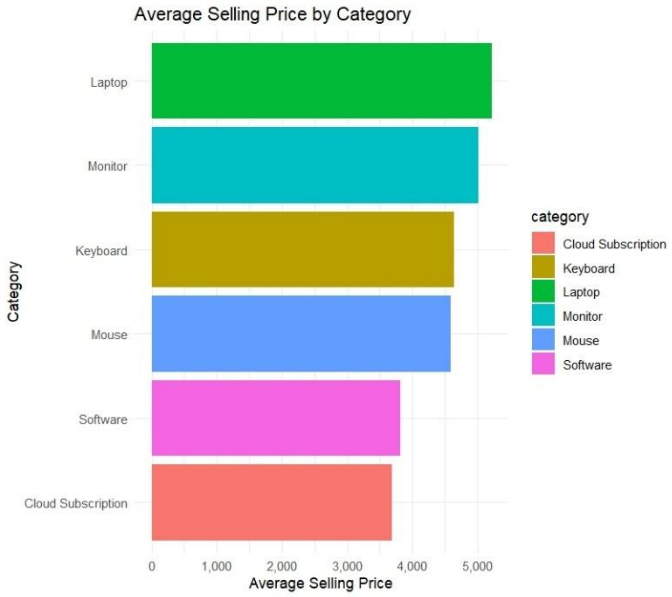


Products data

High-markup categories serve as key drivers of overall profitability, highlighting the importance of maintaining focus on these product lines. Certain high-volume items appear to be underpriced, suggesting that moderate price adjustments could improve margins without significantly affecting demand. The software category emerges as the most frequent and competitively priced segment, reinforcing its role as a strong contributor within the product portfolio.







Suggestions

To use more focused marketing tactics, the business should divide up its clientele by age and wealth. To guarantee product availability during times of strong demand, inventory levels should be modified to correspond with seasonal sales peaks. Reducing picking and delivery times will improve logistics overall and increase customer satisfaction. To increase profitability, pricing for products with low margins but strong sales should also be reevaluated. Lastly, increasing the brand's visibility in big cities would aid in gaining new clients and growing market share.

Part 3

Overall SPC Summary (All Products)

Table 1: Summary of SPC Statistics for All Product Types

	Product	Total_S_Above3Sigma	S_First3	S_Last3	Longest_S_Between1Sigma	X_Seq_GE4	X_First3	X_Last3
CLO	CLO	0			35	14	122, 179, 192	557, 604, 628
KEY	KEY	0			15	27	99, 112, 172	687, 698, 726
LAP	LAP	0			19	11	119, 129, 153	348, 359, 374
MON	MON	0			34	22	134, 171, 179	566, 610, 615
MOU	MOU	1	592	592	16	25	194, 233, 249	768, 777, 807
SOF	SOF	0			21	27	133, 202, 237	774, 803, 842

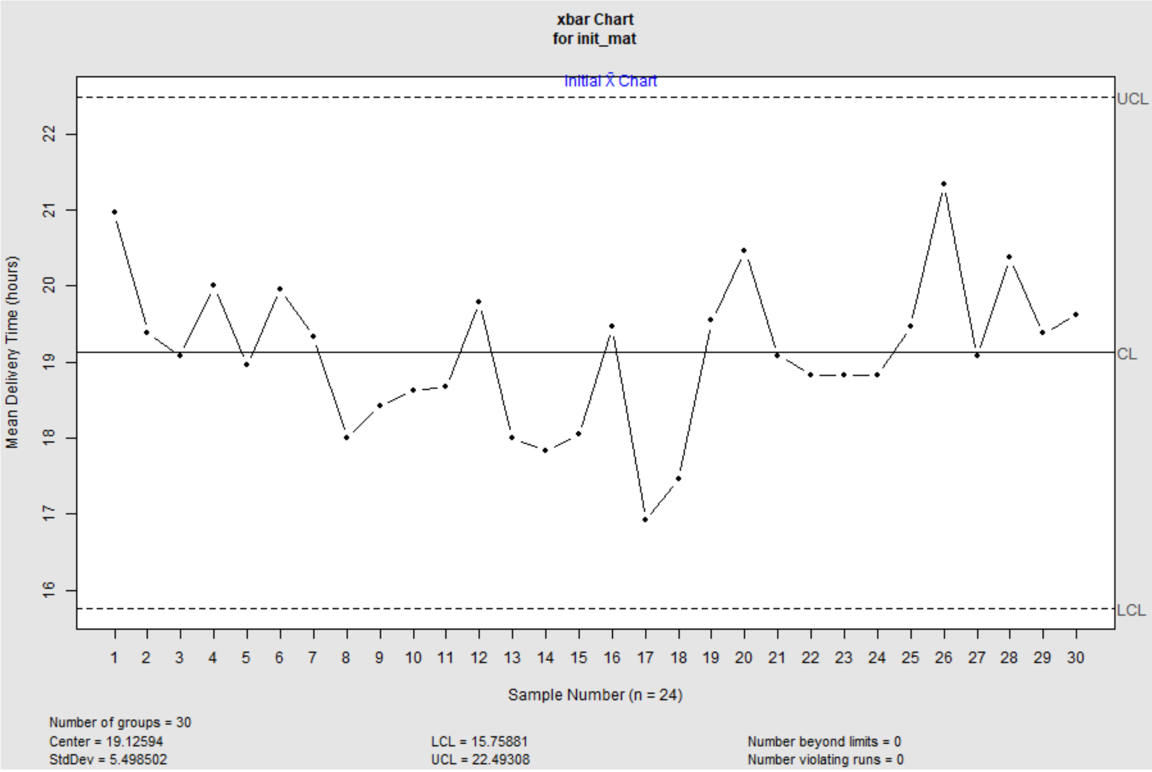
Product: CLO

Process Capability Summary

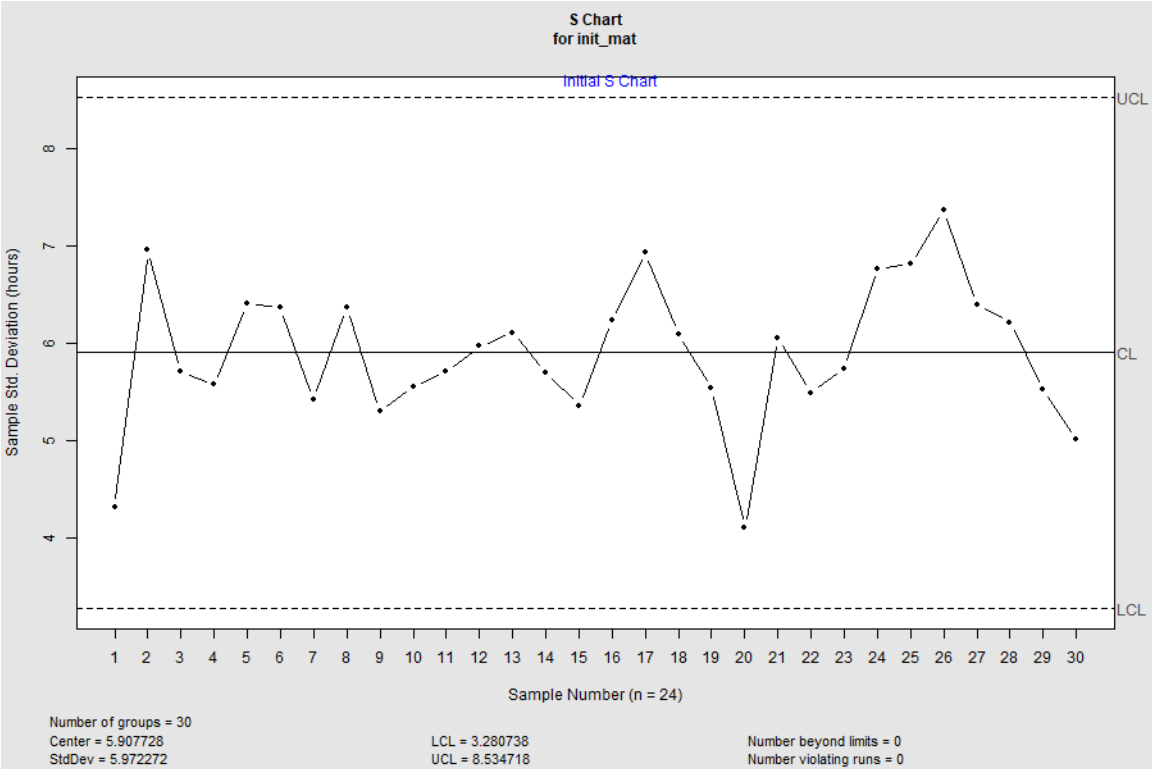
Process Capability Indices for CLO

Cp	Cpl	Cpu	Cpk	capable
0.8977458	1.078754	0.7167378	0.7167378	false

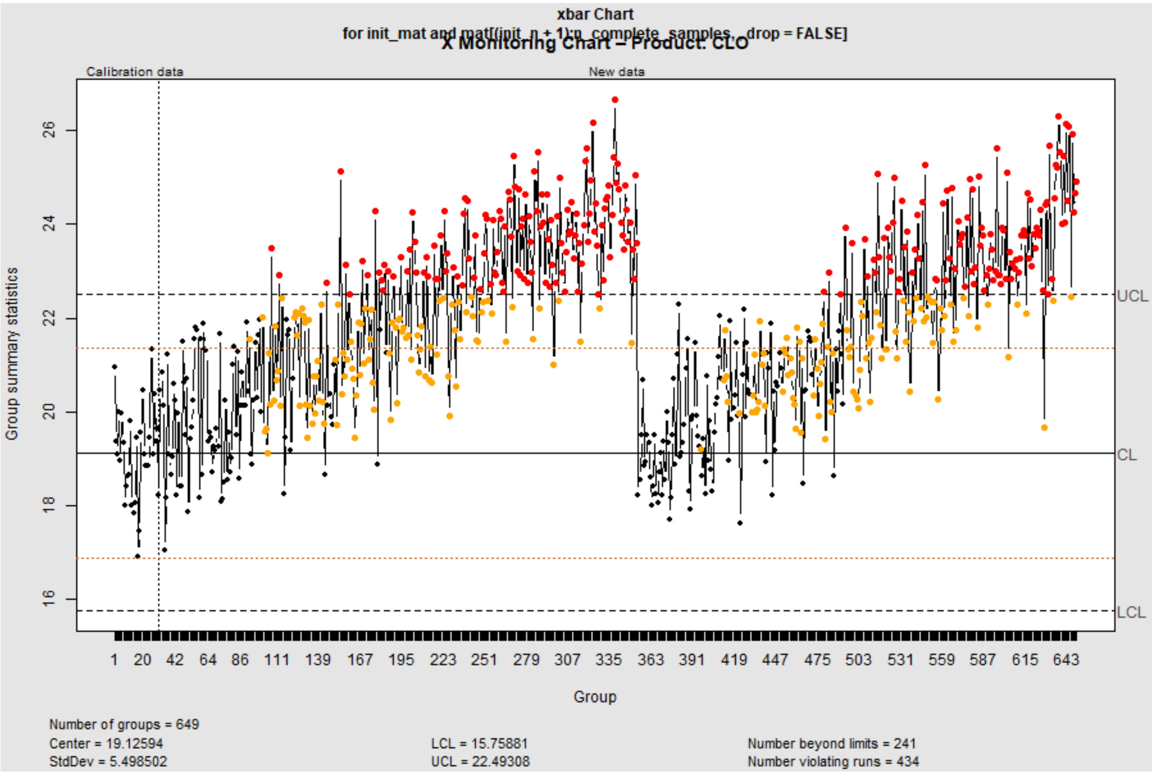
X̄ (Mean) Control Chart – Initial 30 Samples



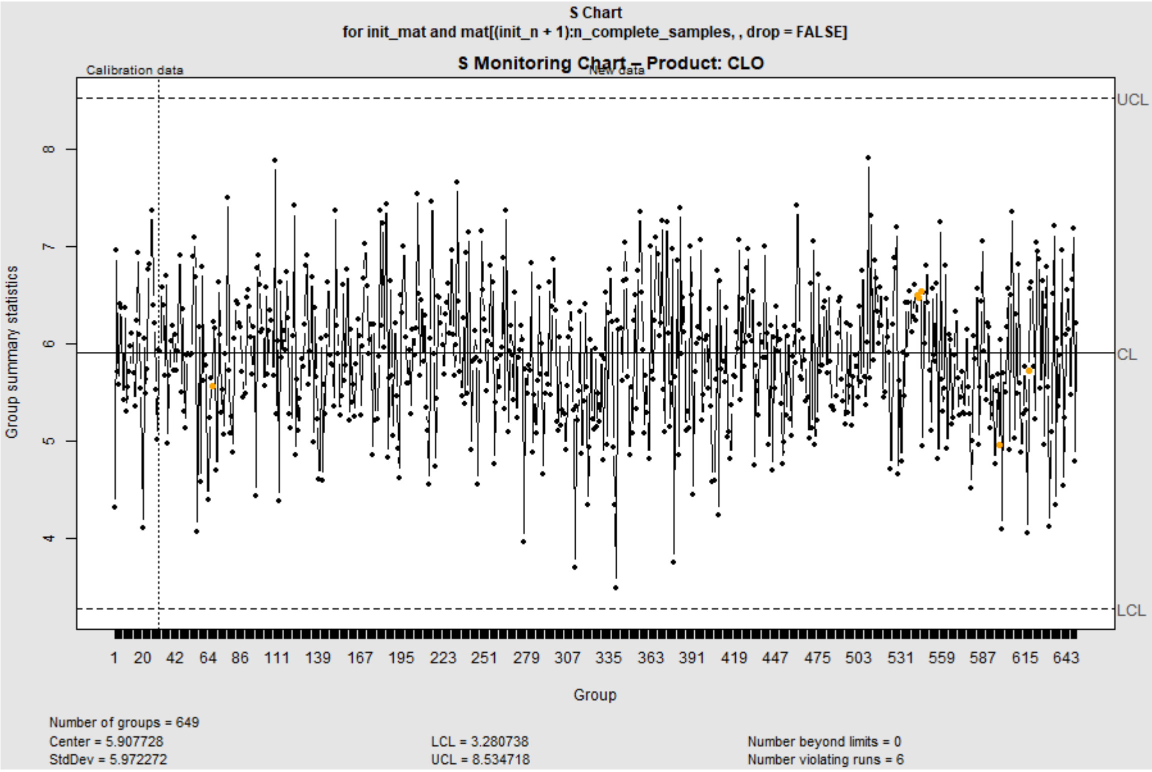
S (Standard Deviation) Control Chart – Initial 30 Samples



X̄ (Mean) Control Chart – Monitoring Samples (31+)



S (Standard Deviation) Control Chart – Monitoring Samples (31+)



Insights and Interpretation

The process for CLO is not capable (Cpk = 0.717), indicating high variation or misalignment with specifications.

The S-chart indicates stable variability with no samples beyond +3σ limits.

The X̄-chart indicates recurring mean shifts (368 samples outside ±2σ), suggesting potential assignable causes or process drift.

The longest stable run within ±1σ on the S-chart is 35 samples, indicating a consistent process over a long period.

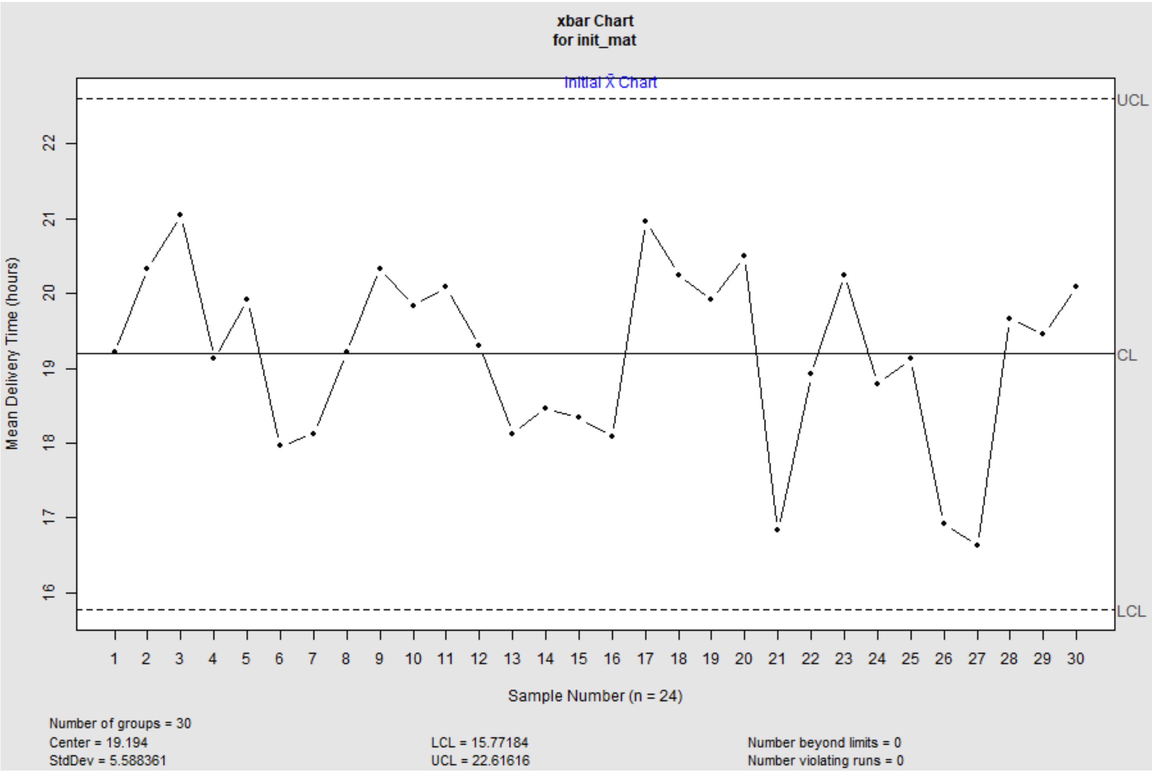
Product: KEY

Process Capability Summary

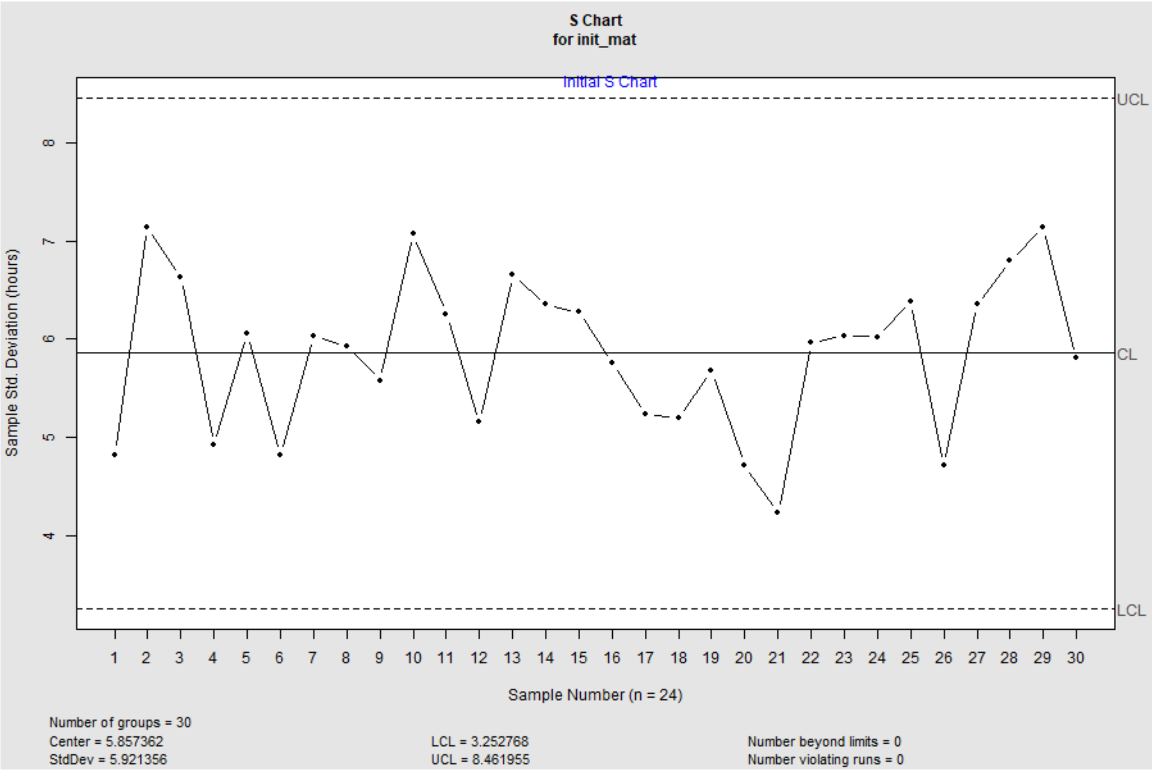
Process Capability Indices for KEY

Cp	Cpl	Cpu	Cpk	capable
0.9171375	1.104921	0.7293536	0.7293536	false

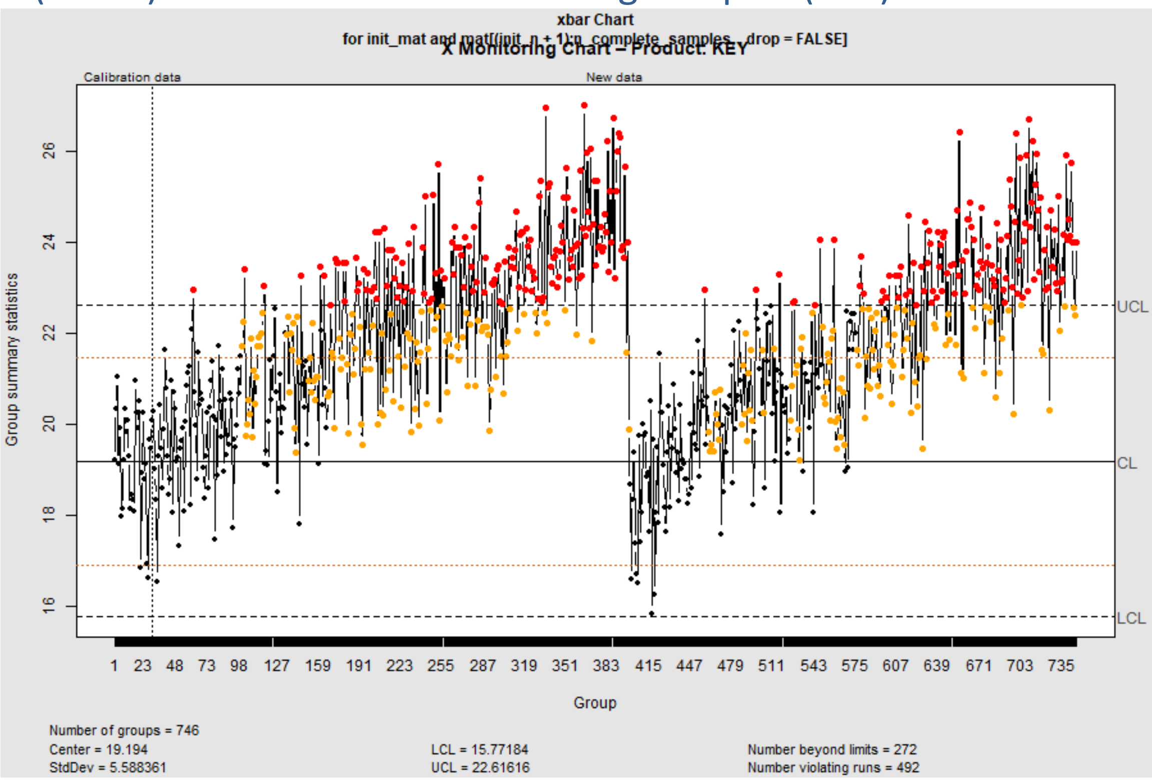
X̄ (Mean) Control Chart – Initial 30 Samples



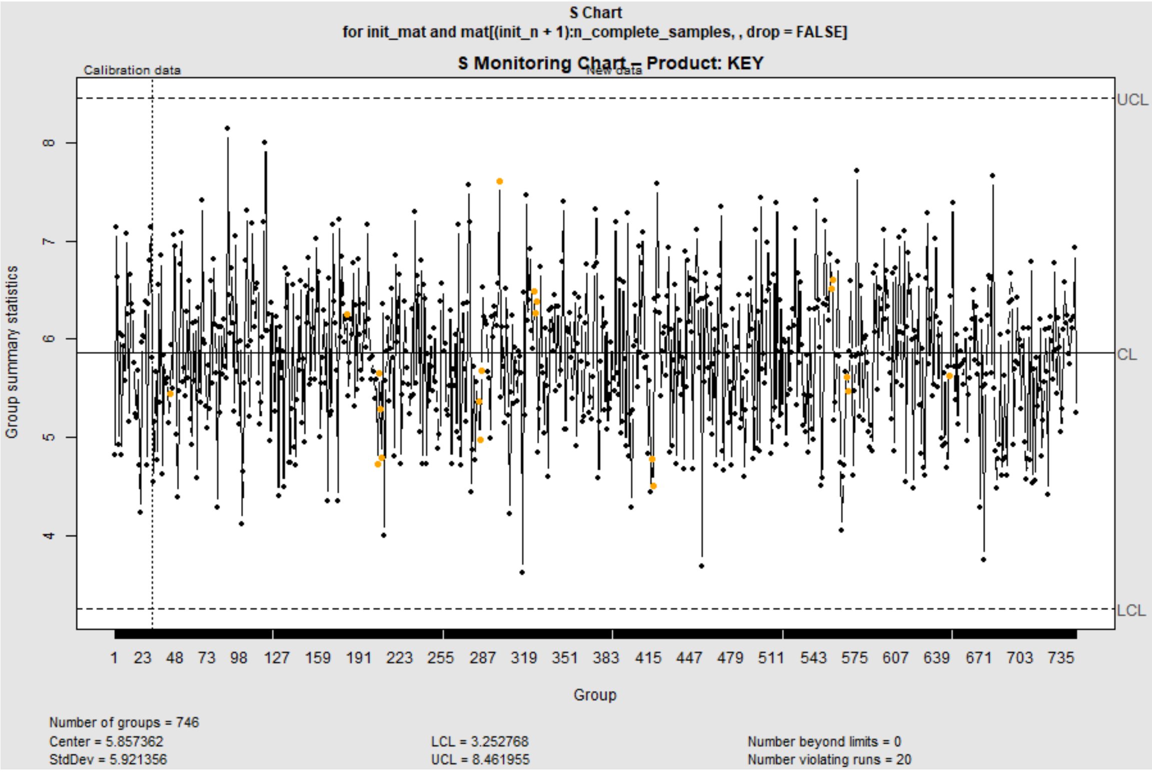
S (Standard Deviation) Control Chart – Initial 30 Samples



X̄ (Mean) Control Chart – Monitoring Samples (31+)



S (Standard Deviation) Control Chart – Monitoring Samples (31+)



Insights and Interpretation

The process for KEY is not capable (Cpk = 0.729), indicating high variation or misalignment with specifications.

The S-chart indicates stable variability with no samples beyond +3σ limits.

The X̄-chart indicates recurring mean shifts (417 samples outside ±2σ), suggesting potential assignable causes or process drift.

The longest stable run within ±1σ on the S-chart is 15 samples, indicating a consistent process over a long period.

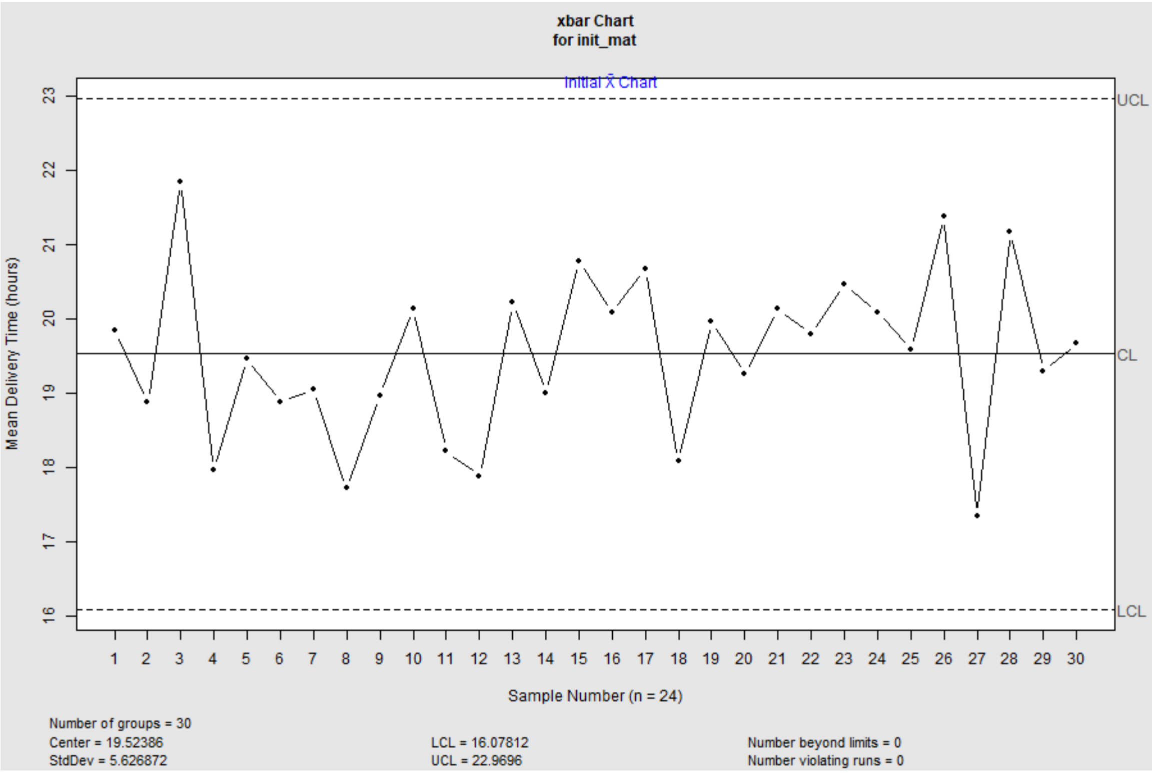
Product: LAP

Process Capability Summary

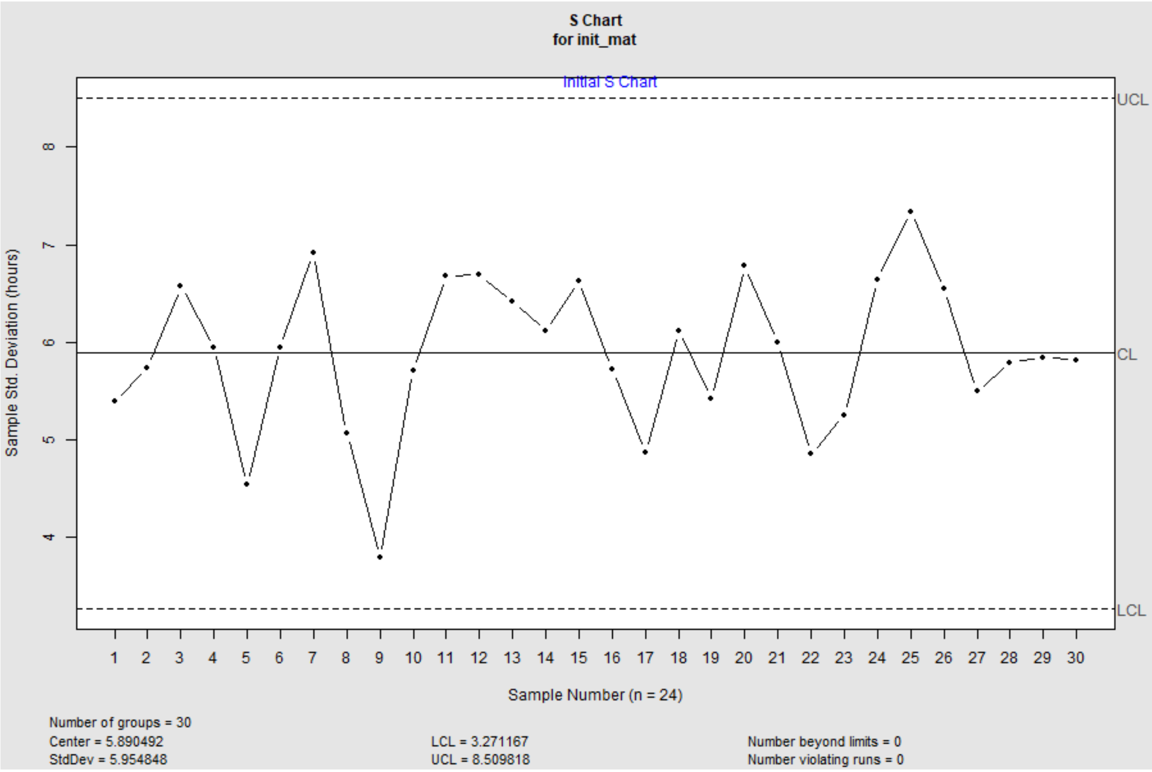
Process Capability Indices for LAP

Cp	Cpl	Cpu	Cpk	capable
0.8987816	1.101345	0.6962187	0.6962187	false

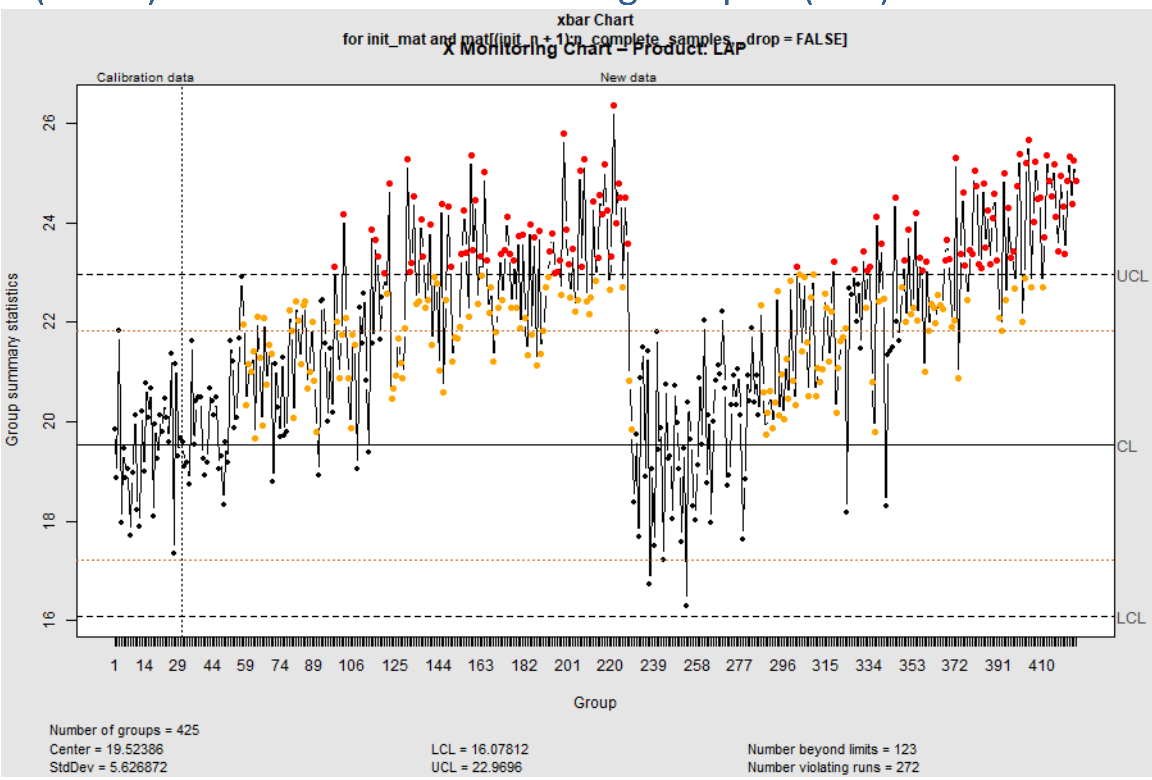
X̄ (Mean) Control Chart – Initial 30 Samples



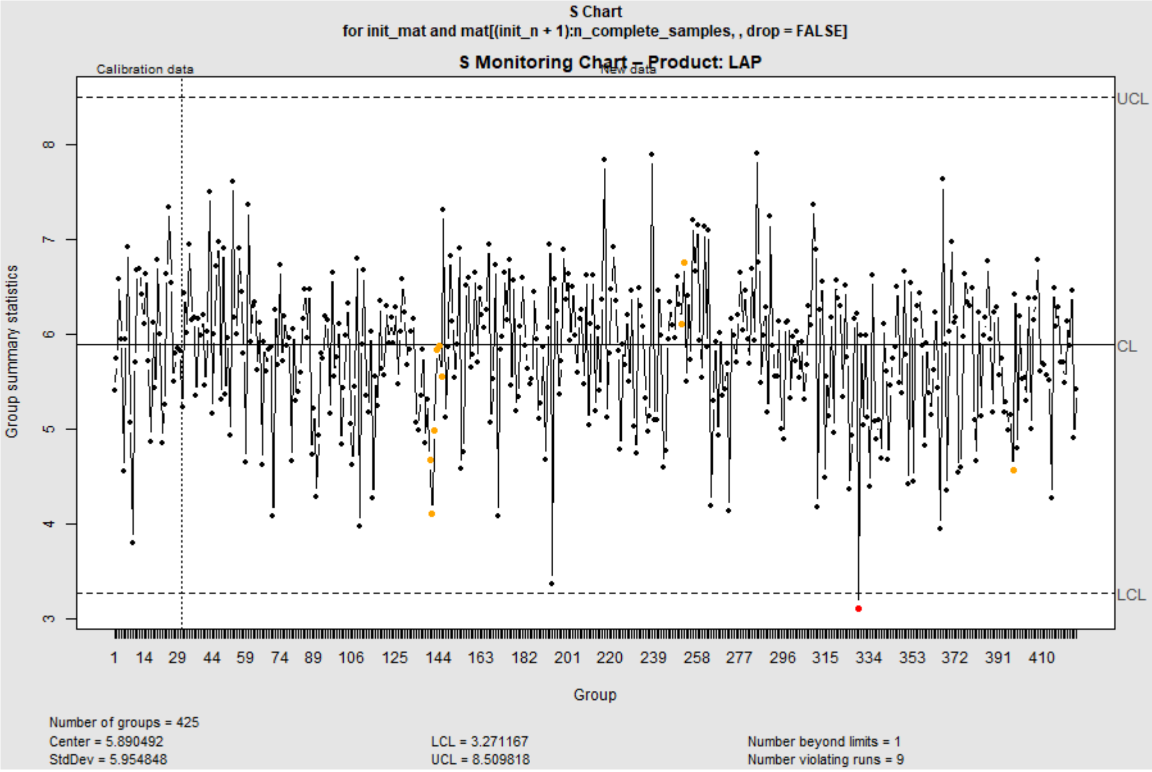
S (Standard Deviation) Control Chart – Initial 30 Samples



X̄ (Mean) Control Chart – Monitoring Samples (31+)



S (Standard Deviation) Control Chart – Monitoring Samples (31+)



Insights and Interpretation

The process for LAP is not capable (Cpk = 0.696), indicating high variation or misalignment with specifications.

The S-chart indicates stable variability with no samples beyond +3σ limits.

The X̄-chart indicates recurring mean shifts (224 samples outside ±2σ), suggesting potential assignable causes or process drift.

The longest stable run within ±1σ on the S-chart is 19 samples, indicating a consistent process over a long period.

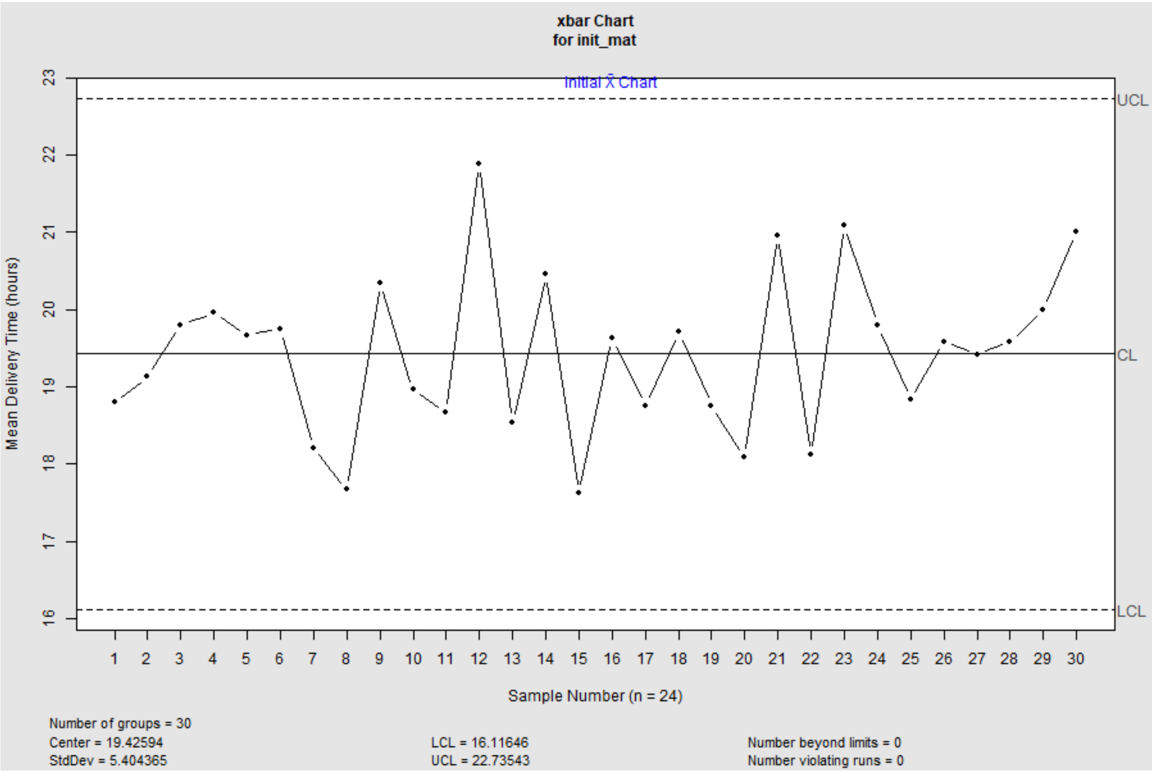
Product: MON

Process Capability Summary

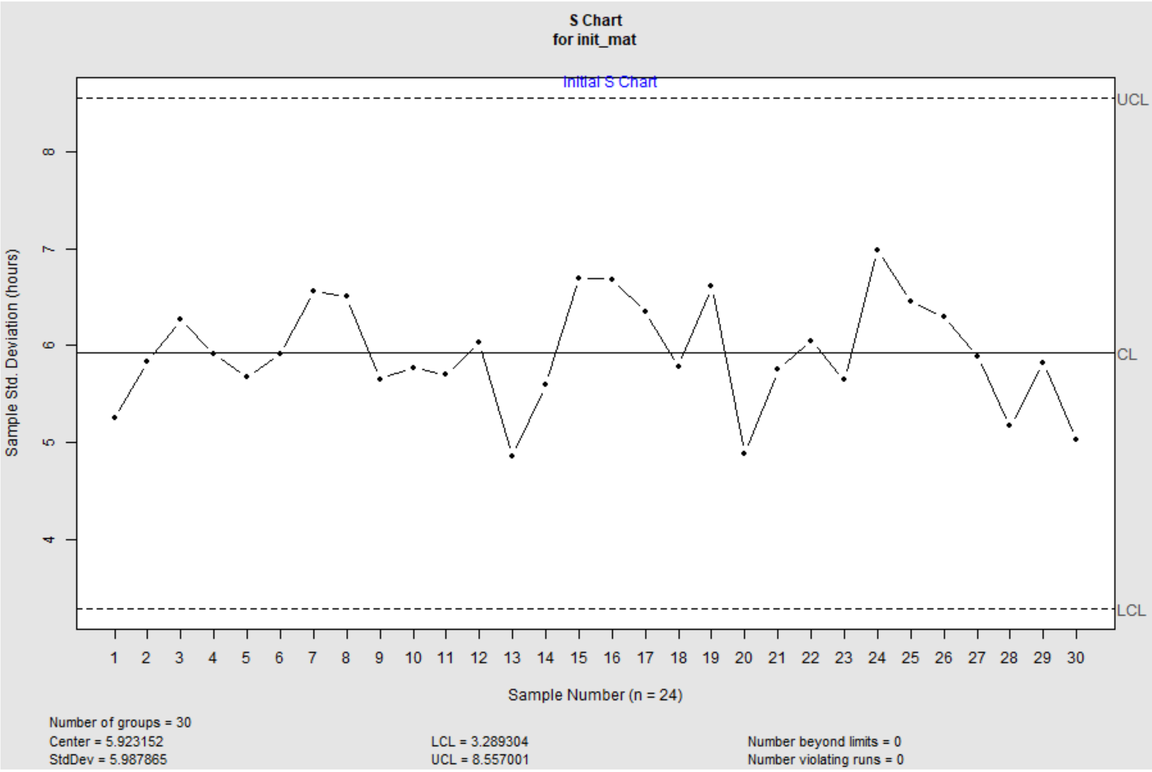
Process Capability Indices for MON

Cp	Cpl	Cpu	Cpk	capable
0.889049	1.078528	0.6995705	0.6995705	false

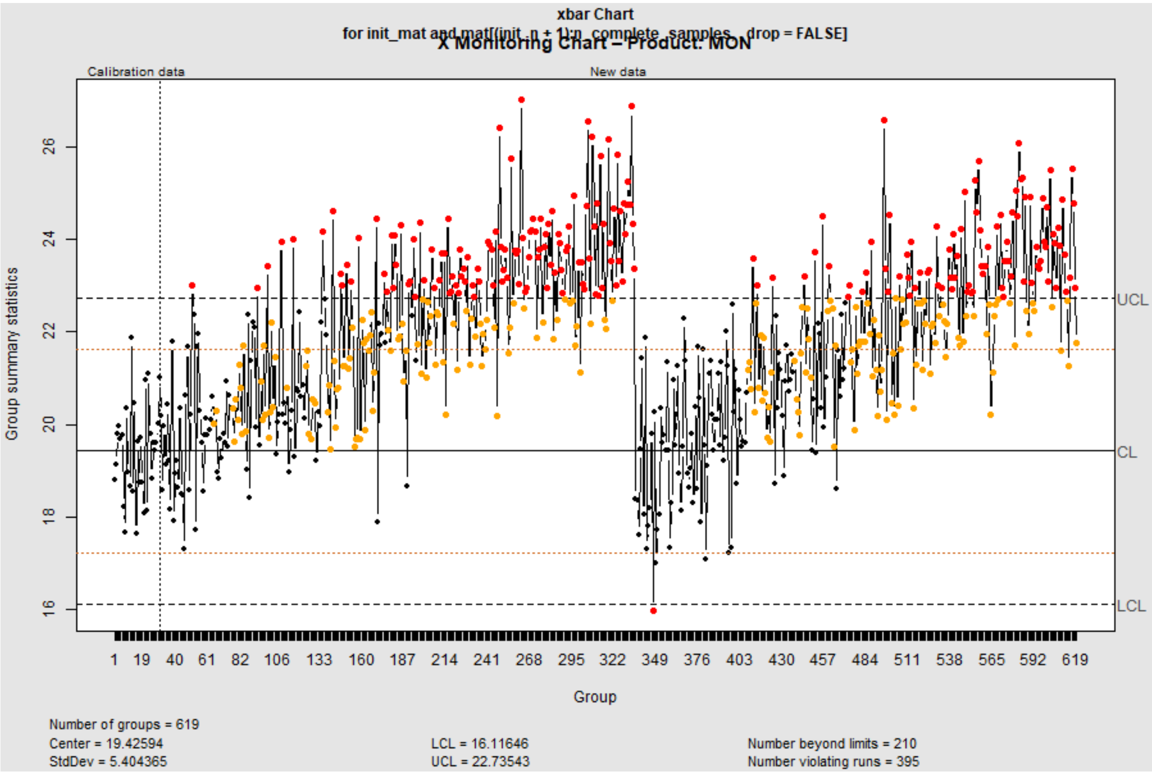
X̄ (Mean) Control Chart – Initial 30 Samples



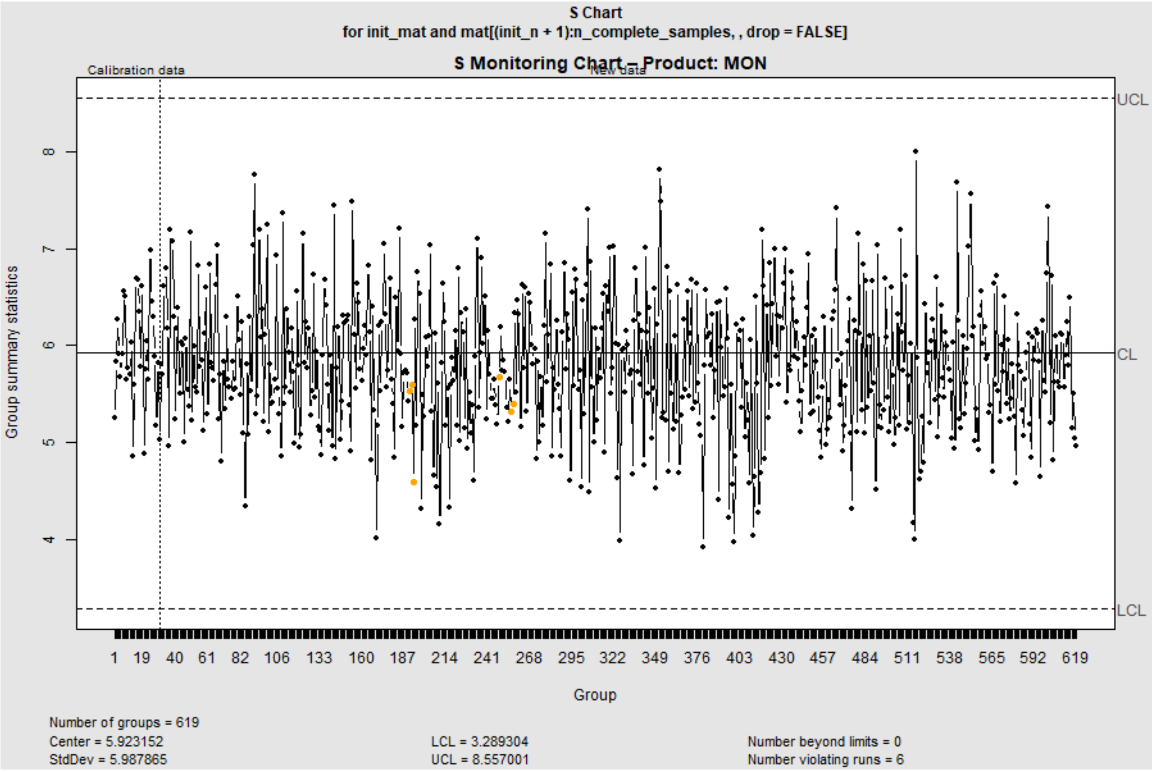
S (Standard Deviation) Control Chart – Initial 30 Samples



X̄ (Mean) Control Chart – Monitoring Samples (31+)



S (Standard Deviation) Control Chart – Monitoring Samples (31+)



Insights and Interpretation

The process for MON is not capable (Cpk = 0.7), indicating high variation or misalignment with specifications.

The S-chart indicates stable variability with no samples beyond +3σ limits.

The X̄-chart indicates recurring mean shifts (339 samples outside ±2σ), suggesting potential assignable causes or process drift.

The longest stable run within ±1σ on the S-chart is 34 samples, indicating a consistent process over a long period.

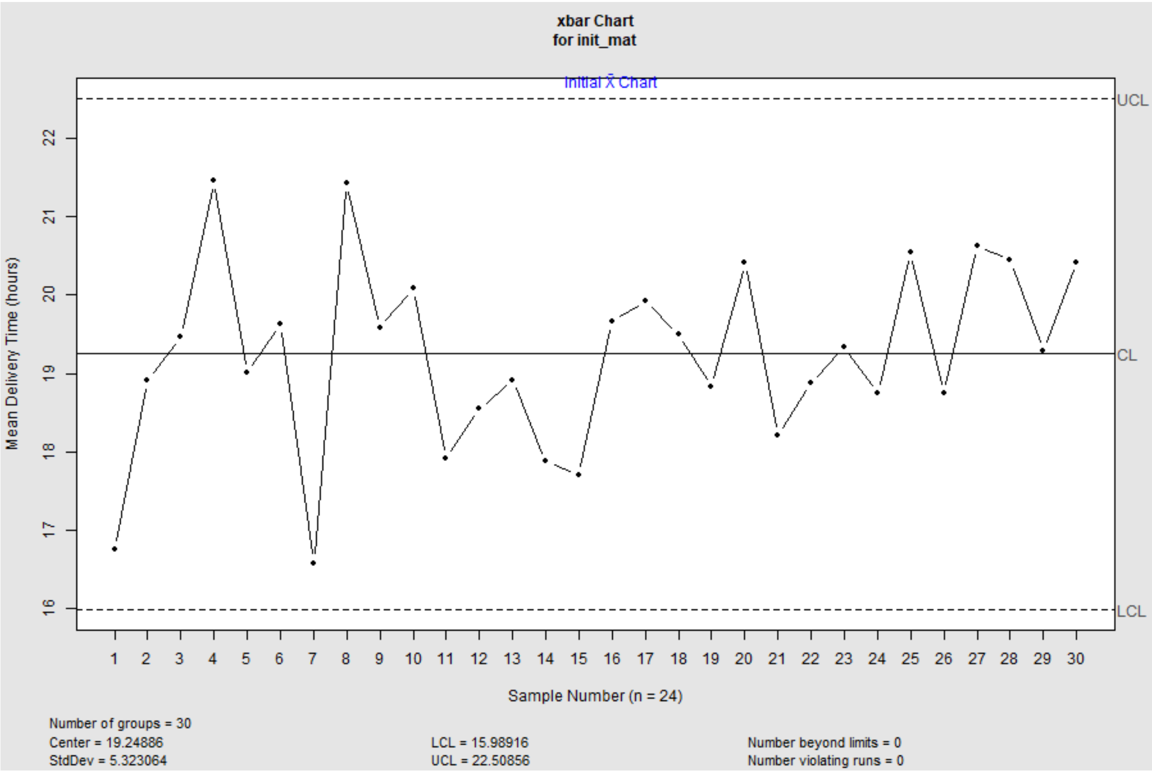
Product: MOU

Process Capability Summary

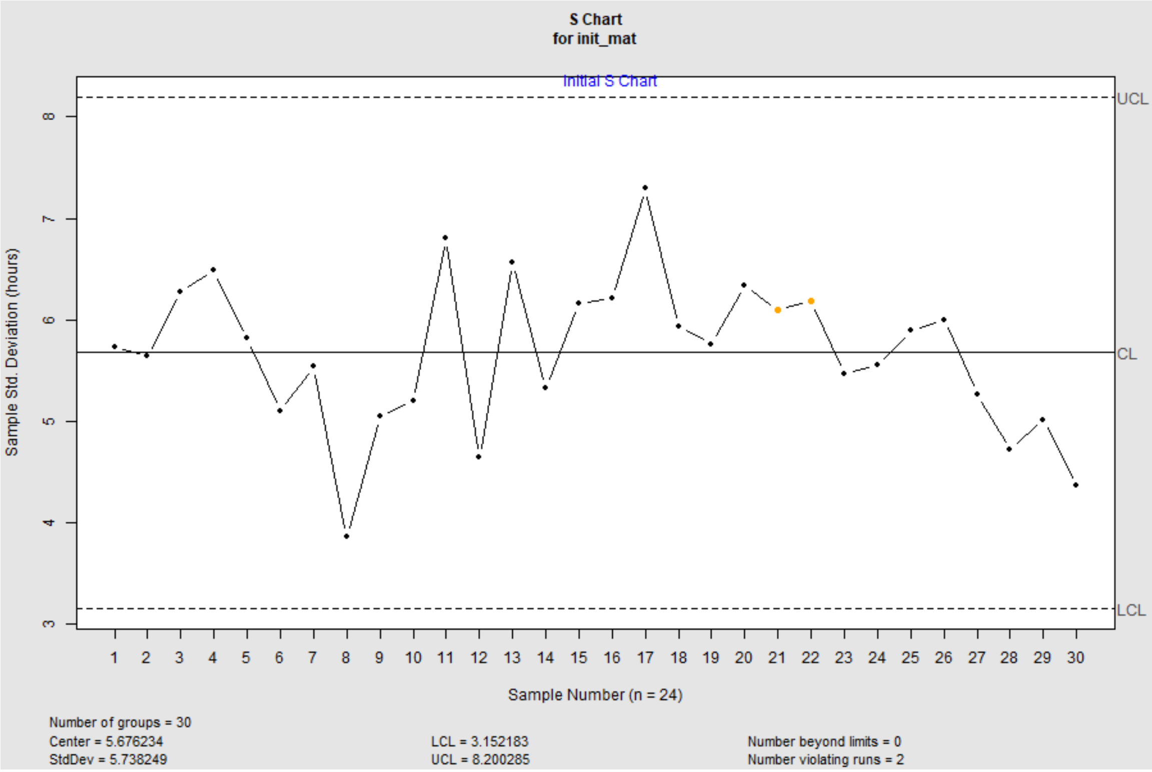
Process Capability Indices for MOU

Cp	Cpl	Cpu	Cpk	capable
0.9151848	1.103799	0.726571	0.726571	false

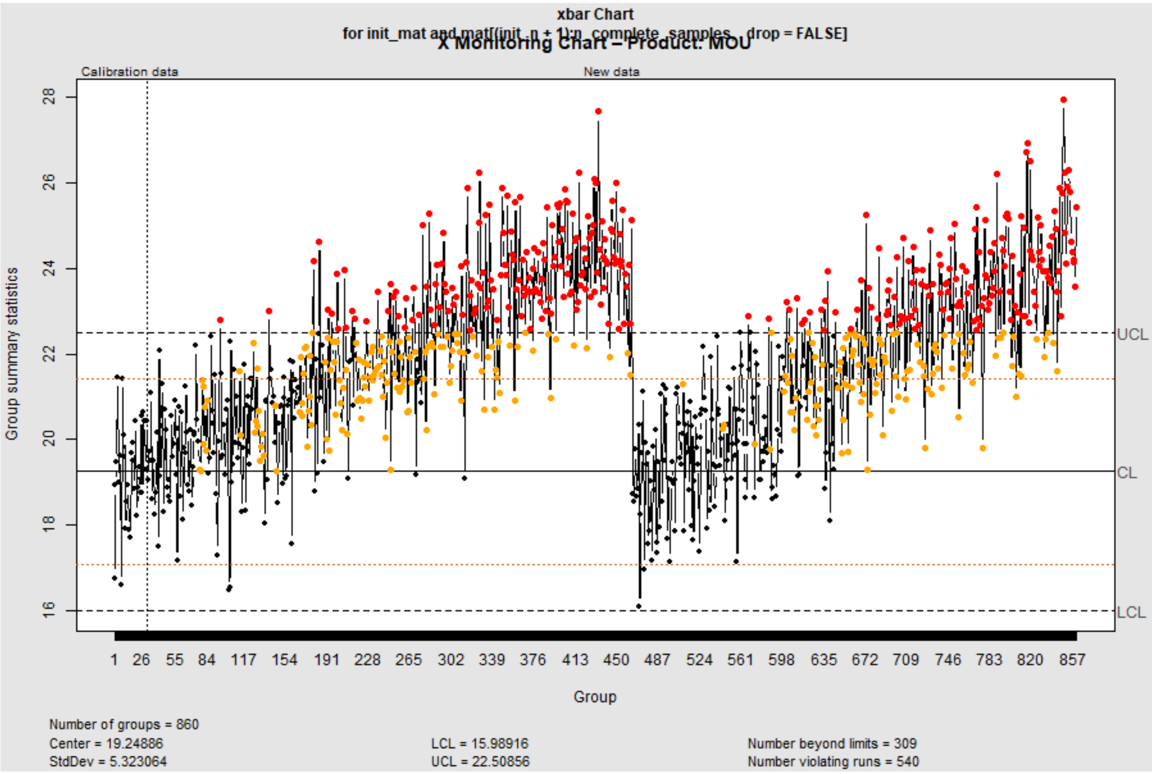
$\bar{X}$ (Mean) Control Chart – Initial 30 Samples



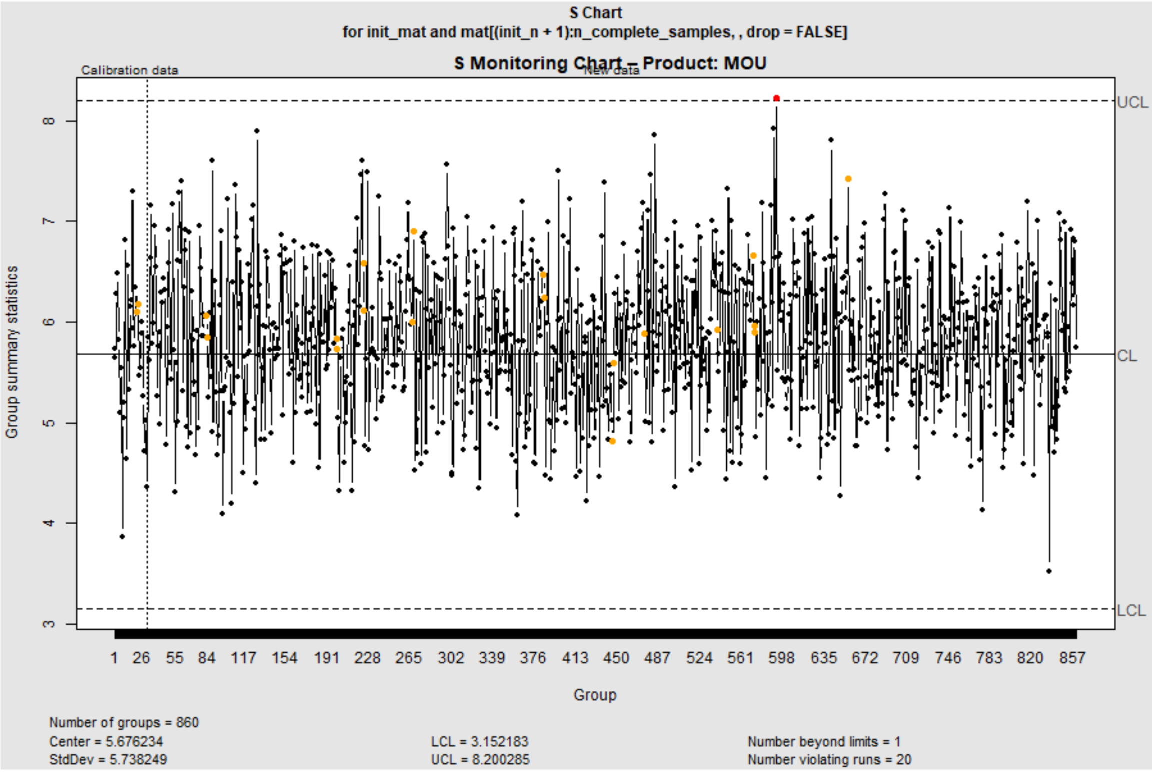
S (Standard Deviation) Control Chart – Initial 30 Samples



X̄ (Mean) Control Chart – Monitoring Samples (31+)



S (Standard Deviation) Control Chart – Monitoring Samples (31+)



Insights and Interpretation

The process for MOU is not capable (Cpk = 0.727), indicating high variation or misalignment with specifications.

The S-chart shows 1 sample exceeding the +3σ limit, suggesting occasional high variability.

The X̄-chart indicates recurring mean shifts (491 samples outside ±2σ), suggesting potential assignable causes or process drift.

The longest stable run within ±1σ on the S-chart is 16 samples, indicating a consistent process over a long period.

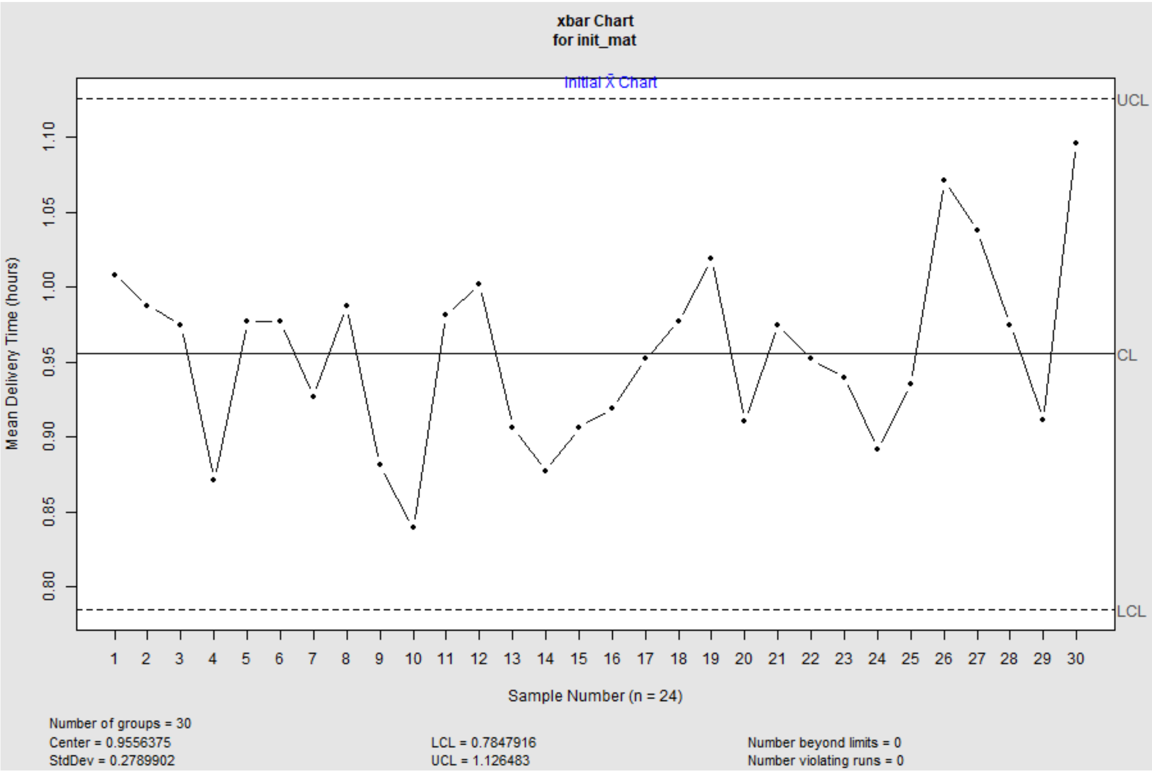
Product: SOF

Process Capability Summary

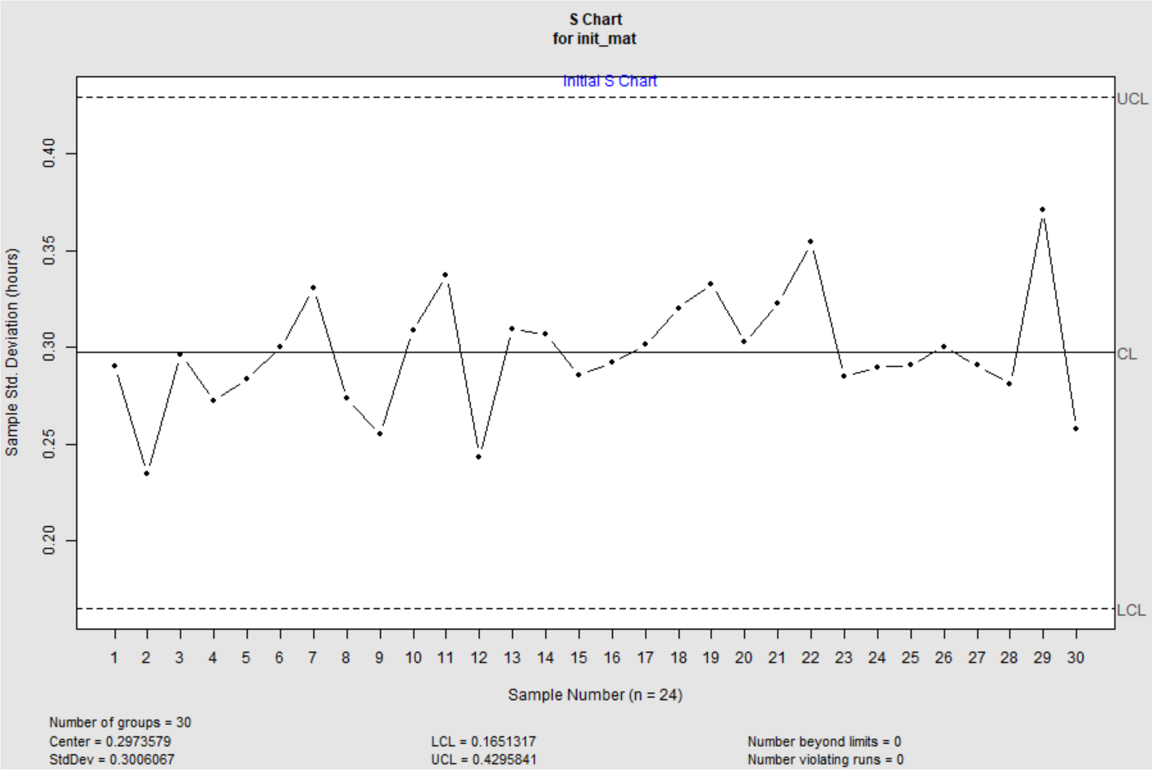
Process Capability Indices for SOF

Cp	Cpl	Cpu	Cpk	capable
18.13524	1.082872	35.1876	1.082872	false

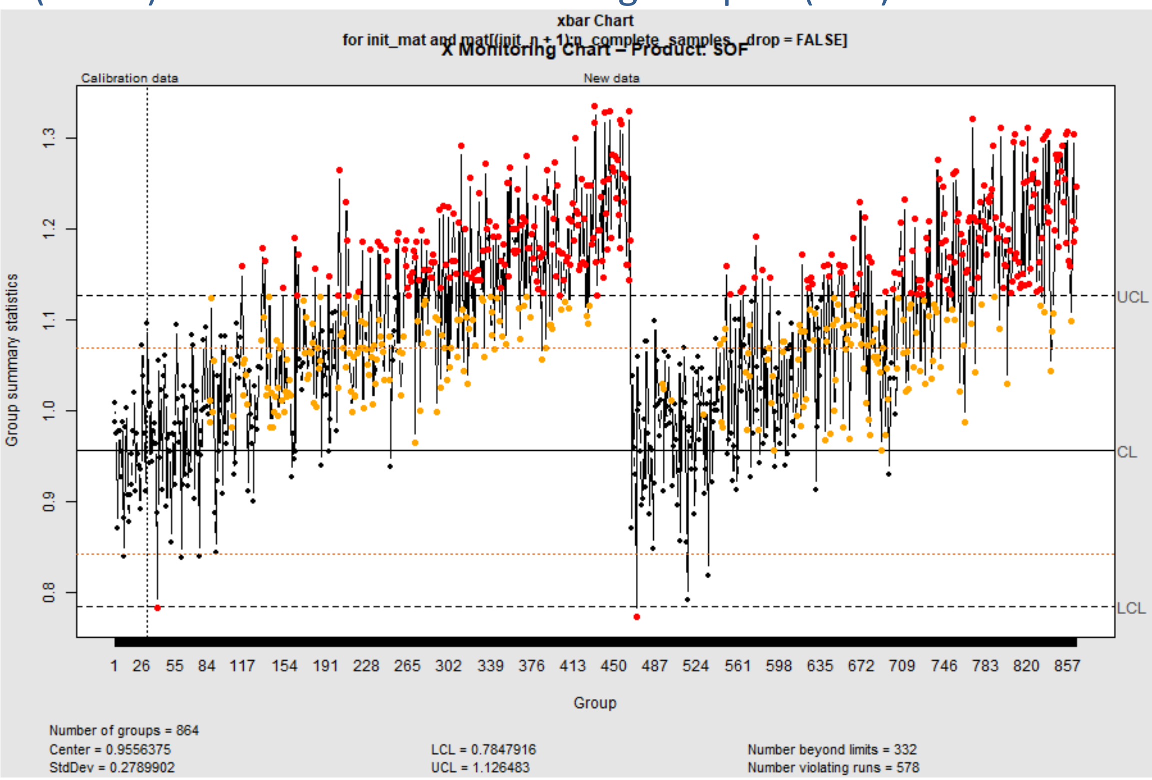
X̄ (Mean) Control Chart – Initial 30 Samples



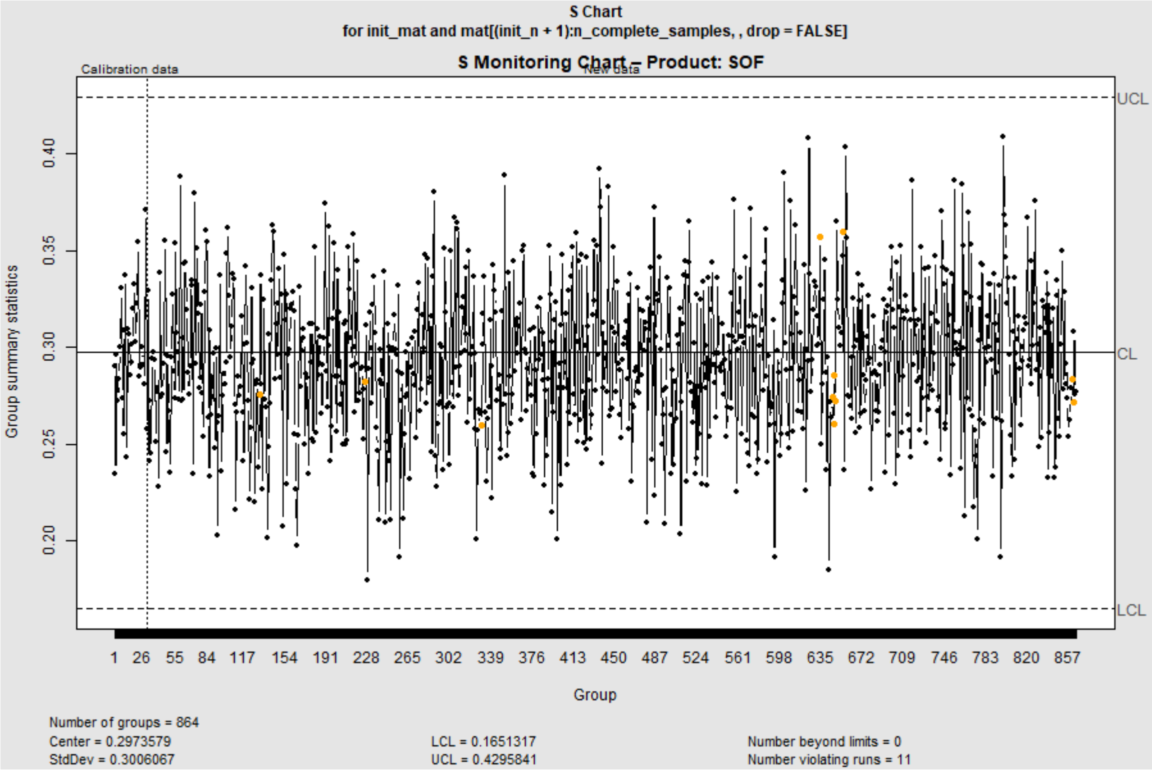
S (Standard Deviation) Control Chart – Initial 30 Samples



X̄ (Mean) Control Chart – Monitoring Samples (31+)



S (Standard Deviation) Control Chart – Monitoring Samples (31+)



Insights and Interpretation

The process for SOF is marginally capable (Cpk = 1.083), which may lead to some out-of-spec deliveries.

The S-chart indicates stable variability with no samples beyond +3σ limits.

The X̄-chart indicates recurring mean shifts (487 samples outside ±2σ), suggesting potential assignable causes or process drift.

The longest stable run within ±1σ on the S-chart is 21 samples, indicating a consistent process over a long period.

Overall Capability Summary

Table: Ranked Process Capability (Highest to Lowest Cpk)

product	Cpk	capable
SOF	1.0828720	false
KEY	0.7293536	false
MOU	0.7265710	false
CLO	0.7167378	false
MON	0.6995705	false
LAP	0.6962187	false

0 out of 6 product types achieved Cpk ≥ 1.3, indicating capable and controlled processes.

Part4

4.1 Type 1 error

This section provides the theoretical Type I (Manufacturer's) Error probabilities for the standard Statistical Process Control rules. These assume a stable, normally distributed process with independent samples.

Rule A — One Point Beyond ±3σ Control Limits
A single sample point that falls outside the Upper Control Limit (UCL) or Lower Control Limit (LCL), set at ±3σ from the process mean. This is the same A as mentioned in 3.4.

Type I Error Probability (α):

$$\alpha_A = 2 \cdot \big(1 - \Phi(3)\big) \approx 0.0027$$

$$\alpha \approx 0.0027 \Rightarrow 0.27\%$$

Rule B — Two Out of Three Consecutive Points Beyond ±2σ on the Same Side
Within three consecutive samples, at least two points fall beyond the ±2σ region on the same side of the centerline.

Type I Error Probability for (α):

$$> +2\sigma = 3p^2(1-p) + p^3 = 0.001529156 \text{ where } p = 1 - \Phi(2).$$

$$\alpha = 2(0.001529156), \text{both sides}$$

$$\alpha \approx 0.00306 \Rightarrow 0.306\%$$

This links with 3.4 C values that require 4 consecutive points outside the second upper control limit.

$$(1 \text{ sample} > +2\sigma) = 0.0228, \text{ so } \alpha = (0.0228)^4 \approx 2.7 \times 10^{-7} \rightarrow 0.000027\%$$

Rule C — Seven Consecutive Points on the Same Side of the Centreline
Seven consecutive sample points fall either all above or all below the centreline. In 3.4 they just ask to calculate the most consecutive points so no exact number of points thus no theoretical probability.

Type I Error Probability (α):

$$\alpha = 0.5^7 = 0.0078125 \Rightarrow 0.78\%$$

$$\alpha = 2 \cdot 0.5^7 = 0.015625 \Rightarrow 1.56\%, \text{both sides}$$

Rule	Condition	Type I Error (α)
A	One point beyond ±3σ	0.27%
B	2 out of 3 beyond ±2σ (same side)	0.306%
C	7 consecutive on same side	0.78% (one side) / 1.56% (either side)

Given

4.2 Type 2 error

CL for $\bar{X}$-chart: 25.05

UCL=25.089, LCL=25.011

True mean μ = 25.028

$\sigma_{\bar{X}}$ =0.017 (was 0.013 before shift)

Type II error (β) = probability we fail to detect the shift (Probability $\bar{X}$ is between LCL and UCL given the shifted mean and std)

Calculation

1. Compute z-scores for the control limits under the new distribution:

   $z_U = \frac{25.089-25.028}{0.017} = \frac{0.061}{0.017} = \frac{61}{17} = 3.588235294117647.$

   $z_L = \frac{25.011-25.028}{0.017} = \frac{-0.017}{0.017} = -1.00.$

2. Using the standard normal CDF $\Phi(\cdot)$

   $\Phi(z_U) \approx \Phi(3.588235) \approx 0.999828$

   $\Phi(z_L) = \Phi(-1.00) = 0.158655.$

3. Probability the sample means falls between LCL and UCL (Type II error):

   $$\beta \;=\; \Phi(z_U) - \Phi(z_L) \;\approx\; 0.999828 - 0.158655 \;=\; 0.841173$$

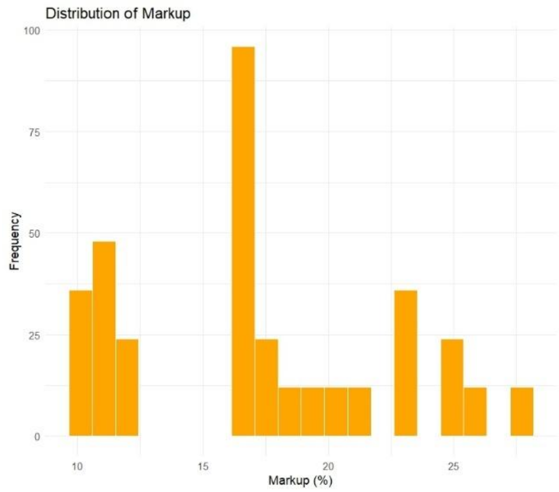
   Thus β =84.12%

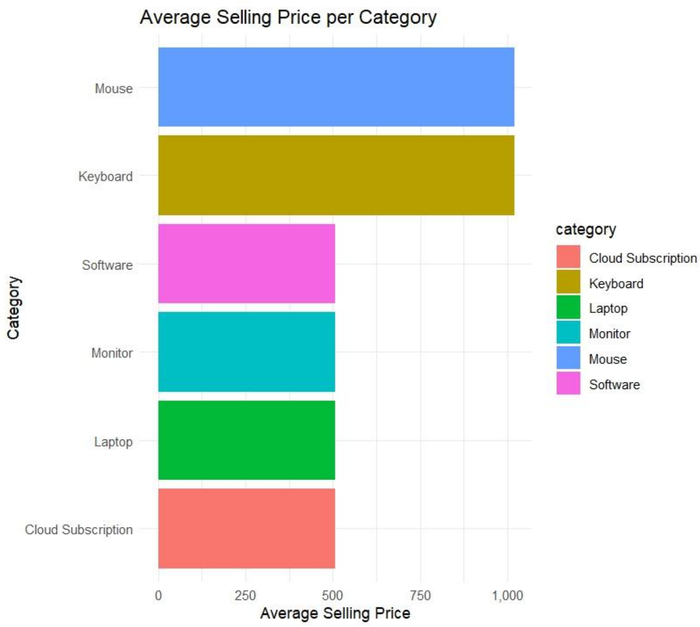
4. Power of the chart (probability of detecting the shift) is $1-\beta$:

   $$\text{Power} = 1-\beta \approx 1 - 0.841173 = 0.158827 \Rightarrow 15.88\%$$

Part 4.3 updated data analysis

I redid the analysis with new updated data. Some tables and data are the same and won't be included as it would be redundant. I am focusing on the change in products head office as it has all the sales and will give good understanding of the change in the data.





As can be seen above the data has changed quite a bit from the original analysis. The average selling price per category has changed and now shows that Mouse and keyboard are the categories with the highest average selling price. The markup and selling price distribution has also changed dramatically. With the markups not being as widely spread throughout. Now rather appearing in batches of different ranges

Part 5 Baristas and time to serve

Summary by Number of Baristas shop 1

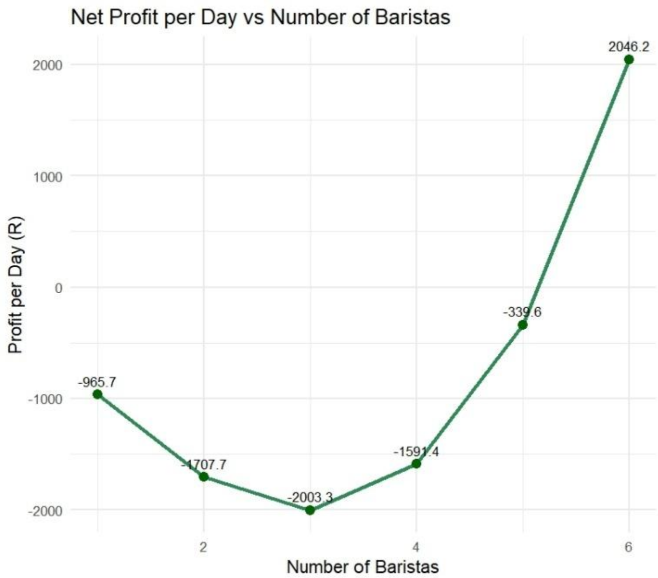
Table 1: Aggregated Summary by Barista Staffing Level

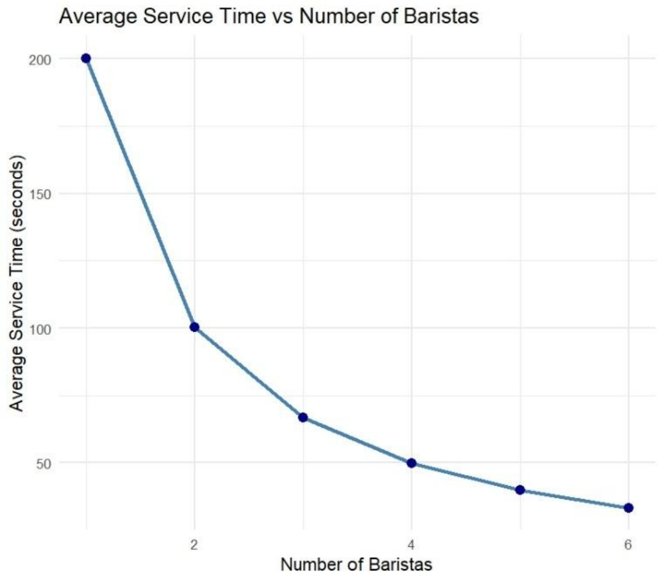
	Baristas	TotalCustomers	CustomersPerDay	MeanSec	MedianSec	Pct_le_thresh1	Pct_le_thresh2	ProfitPerDay
1	1	417	1.142466	200.15588	200	0.00000	0.9592326	-965.7260
2	2	3556	9.742466	100.17098	100	99.71879	100.0000000	-1707.7260
3	3	12126	33.221918	66.61174	67	100.00000	100.0000000	-2003.3425
4	4	29305	80.287671	49.98038	50	100.00000	100.0000000	-1591.3699
5	5	56701	155.345205	39.96183	40	100.00000	100.0000000	-339.6438
6	6	97895	268.205479	33.35565	33	100.00000	100.0000000	2046.1644

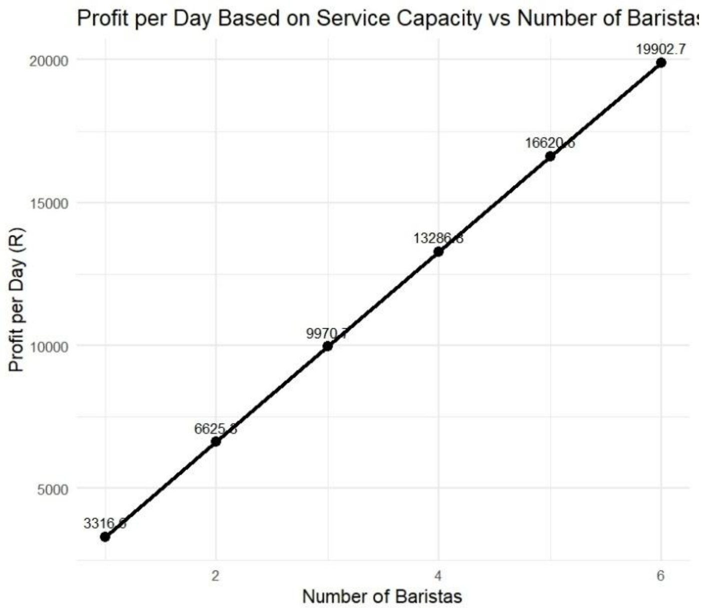
Optimal Staffing Level

Table 2: Optimal Number of Baristas Based on Profit

	Baristas	TotalCustomers	CustomersPerDay	MeanSec	MedianSec	Pct_le_thresh1	Pct_le_thresh2	ProfitPerDay
1	6	97895	268.2055	33.35565	33	100	100	2046.164







Insights and Interpretation

Based on a year's worth of time-to-serve data analysis: 6 is the ideal number of baristas on duty. With 6 people on duty a daily profit of about R2046.16 is anticipated this calculation assumes that every day the same amount of customers come to the coffee shop and not the actual amount. The service time average is 33.4 seconds. In about 120 seconds, 100% of clients are served, and in 180 seconds, 100%. This shows that most clients receive dependable service in three minutes or less, all the while keeping a healthy profit margin. Beyond this threshold, hiring more employees results in diminishing benefits because of increased labour costs each day. Suggestions: 1. Keep six baristas on duty on typical workdays. 2. Keep an eye on service times to make sure dependability remains above 90% within 180 seconds. The third graph also shows that with an increase of number of baristas the profit based on average service time will increase linearly.

Summary by Number of Baristas for shop 2

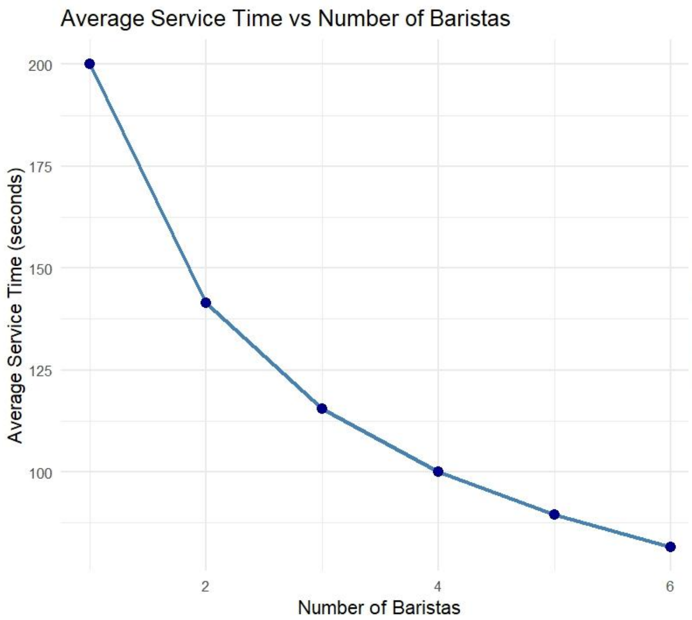
Table 1: Aggregated Summary by Barista Staffing Level

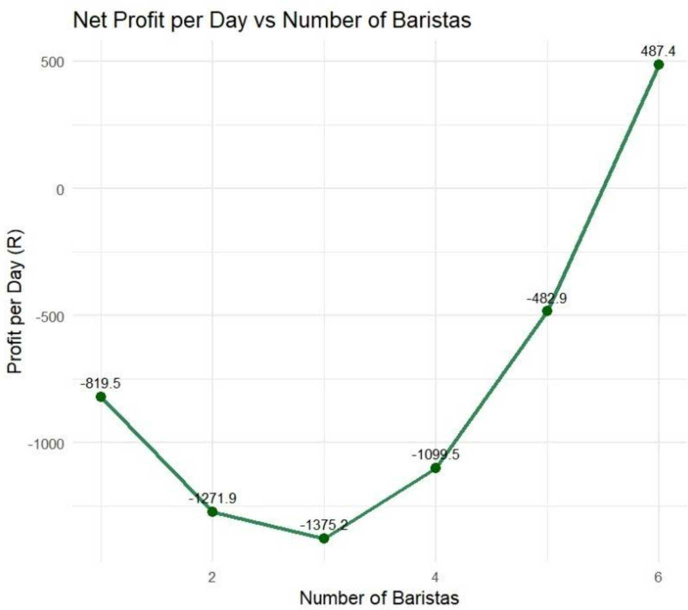
	Baristas	TotalCustomers	CustomersPerDay	MeanSec	MedianSec	Pct_le_thresh1	Pct_le_thresh2	ProfitPerDay
1	1	2196	6.016438	200.16894	200	0.0000000	0.9562842	-819.5068
2	2	8859	24.271233	141.51462	141	0.1241675	100.0000000	-1271.8630
3	3	19768	54.158904	115.44091	116	79.3454067	100.0000000	-1375.2329
4	4	35289	96.682192	100.01527	100	99.9971663	100.0000000	-1099.5342
5	5	54958	150.569863	89.43597	89	100.0000000	100.0000000	-482.9041
6	6	78930	216.246575	81.64272	82	100.0000000	100.0000000	487.3973

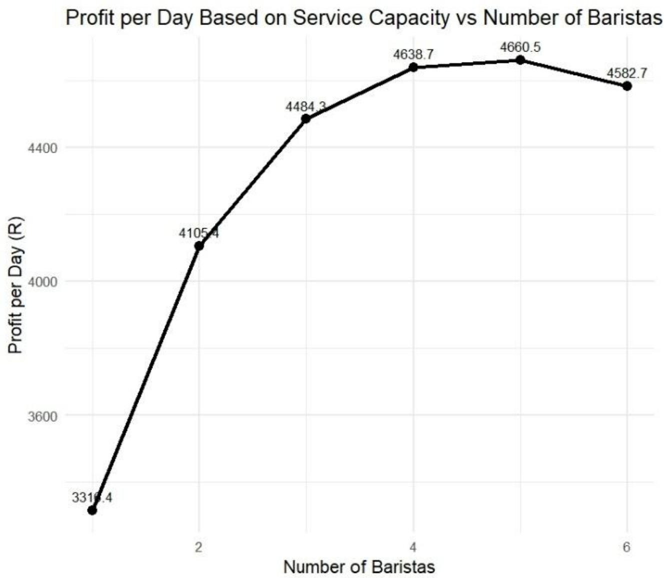
Optimal Staffing Level

Table 2: Optimal Number of Baristas Based on Profit

	Baristas	TotalCustomers	CustomersPerDay	MeanSec	MedianSec	Pct_le_thresh1	Pct_le_thresh2	ProfitPerDay
1	6	78930	216.2466	81.64272	82	100	100	487.3973







Insights and Interpretation

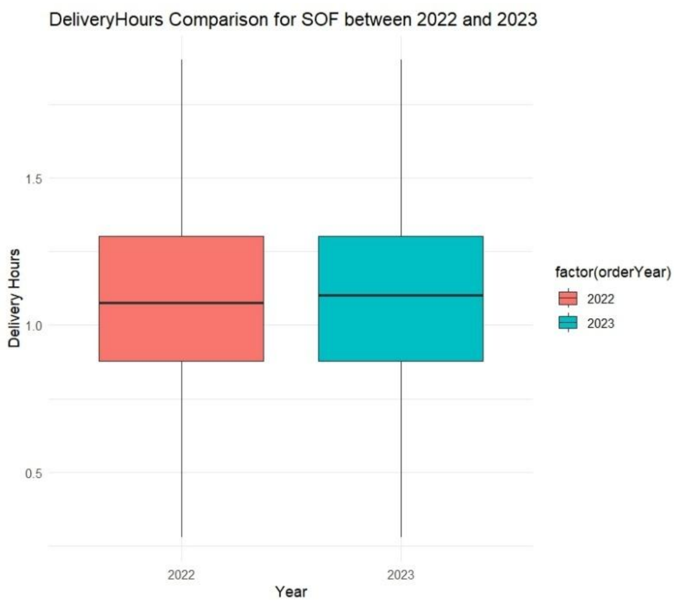
Based on a year's worth of time-to-serve2 data analysis: 6 or 5 is the ideal number of baristas on duty. With 6 baristas on duty a daily net profit of about R487.4 is anticipated this calculation assumes that every day the same amount of customers come to the coffee shop and not the actual amount. The service time average is 81.6 seconds. In about 120 seconds, 100% of clients are served, and in 180 seconds, 100%. This shows that most clients receive dependable service in three minutes or less, all the while keeping a healthy profit margin. Beyond this threshold, hiring more employees results in negative financial benefits because of increased labour costs each day. Suggestions: 1. Keep six baristas on duty on typical workdays. 2. During times of high demand (e.g., weekends, morning rush), add one or two more baristas. 3. Keep an eye on service times to make sure dependability remains above 90% within 180 seconds. The third graph for this data set shows way different outcome then the first data set. Here with the increase of number of baristas the profit increases but starts to plateau and decrease and with this graph it shows that 5 would be the optimal amount of baristas but taking in to account the service time and the net profit per day graph you could argue 6 is better.

Part 6 ANOVA

H_0: There is no significant difference in mean revenue between 2022 and 2023 for product SOF.
H_1: There is a significant difference.

My chosen significance level is 0.05.

	V1	V2	V3	V4	V5	V6
1	Source	SS	DoF	MS	fo	P-value
2	Treatment	0.01	1	0.01	0.13	0.7235637
3	Error	1822.33	19242	0.09	---	---
4	Total	1822.34	19243	---	---	---



A one-factor ANOVA was performed to assess the presence of a significant difference in the mean DeliveryHours for product SOF between the years 2022 and 2023.

The calculated p-value is 0.7236, exceeding the predetermined significance level of $\alpha = 0.05$. Thus, we do not reject the null hypothesis (H_0) and determine that there is no statistically significant difference in mean delivery time for the SOF product line between 2022 and 2023.

The data suggests that the delivery process performance for SOF has maintained stability over the two-year period, showing no significant changes that would need adjustments or corrective measures. This is consistent with SPC principles, indicating that a high p-value points to consistency in the data.

Part 7.1

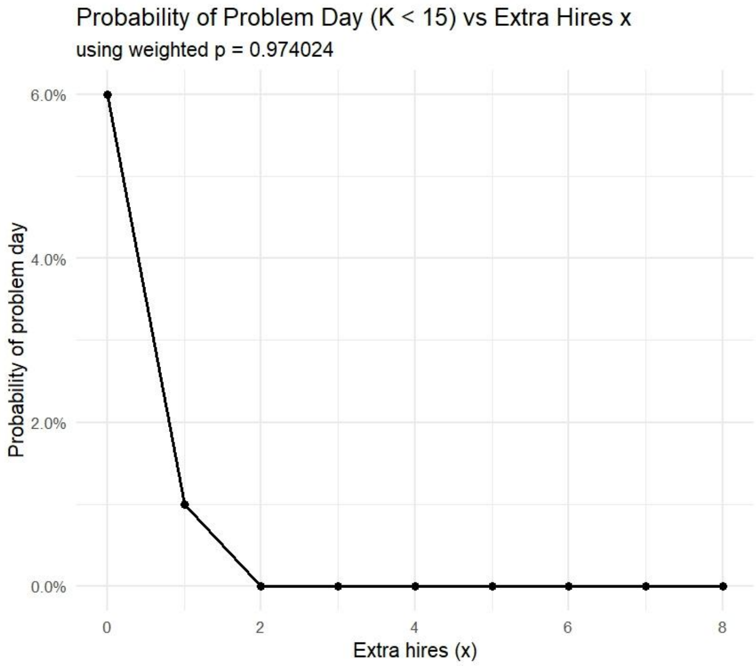
If you set the threshold for reliable service to 15 or more workers in a day the percentage reliability would be:

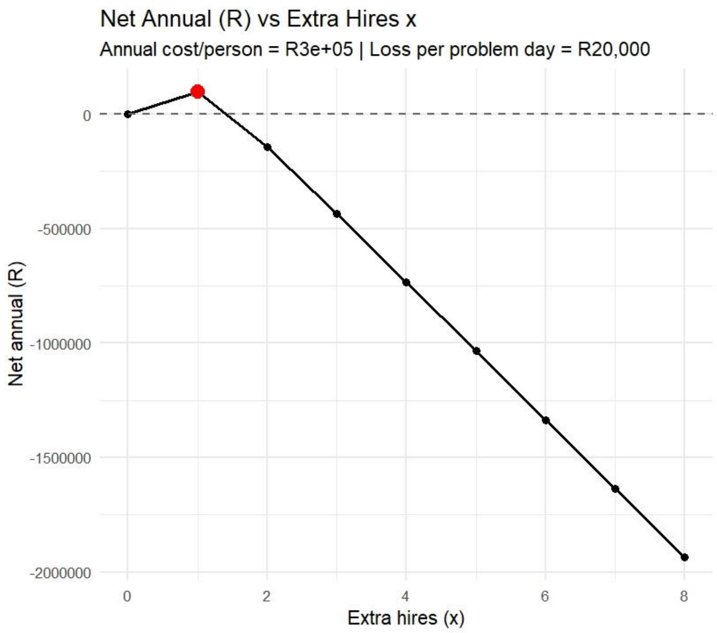
$$\text{Reliable days } (>=15) = 96 + 270 = 366 \text{ days}$$

$$\text{Reliability percentage} = 366 / 397*100 = 92.2\%$$

Considering the frequency distribution of staffing levels and establishing a minimum threshold of 15 or more workers on duty for reliable service, it is estimated that 366 out of 397 observed days (approximately 92.2%) would provide reliable service performance.

Part 7.2





Hiring one extra staff member would be optimal for financial gain for the company: it maximizes net annual profit because it greatly reduces the probability of a problem day (to ~1%), while hiring two or more eliminates almost all remaining risk but the additional costs of hiring more will decrease the net annual profit.

Conclusion

As seen in the analysis did there are always ways to improve the model that a company has. By applying these data-driven actions, the companies can strengthen its market position and achieve sustainable growth.