Project ECSA 2025 – Process Analysis and Quality Improvement

Module: Quality Assurance 344

Student: Reynhardt Bredenkamp

Student No: 26174863

Institution: Stellenbosch University

Date: 23 October 2025

Table of Contents

1. INTRODUCTION

This report presents the analysis and improvement of process performance for a number of different business practices, completed as part of Project ECSA 2025 – Process Analysis and Quality Improvement under the Quality Assurance (QA344) module at Stellenbosch University.

The study applies a data driven approach combine statistical process control, design of experiments, and optimisation modelling to evaluate and enhance the operational performance of the given scenarios presented. Process capability and reliability analyses were used to assess the system stability of processes, while a binomial probability model and profit-optimisation framework was developed to determine the optimal staffing and service reliability levels.

The project demonstrates the practical application of quantitative engineering methods to identify quality control issues, improve efficiency, and maximise profitability in alignment with ECSA Graduate Attribute 4 on problem analysis and performance optimisation.

2 . DATA PREPARATION AND QUALITY REVIEW

2.1 DATA INTEGRITY ASSESSMENT

The first stage of the project involved reviewing all datasets for completeness, consistency and accuracy to ensure they were suitable for accurate quantitative analysis. The datasets included past and future sales data as well as relevant customer and product information. Initial inspection revealed several formatting inconsistencies such as csv exports used semicolon delimiters, header rows were embedded within data, and variable names (e.g., *pickingTime* vs *pickingHours*) differed across files. Duplicate and missing entries were also identified and corrected. A comprehensive DATA Quality Review verified data types, completeness and integrity confirming no significant bias or data loss. All datasets were cleaned and standardised I R using a custom import script that ensured all of the datasets quality issues were addressed. The final integrated datasets provided a consistent, reliable foundation for further analyses.

2.2 DATA STANDARDISATION AND STRUCTURING

After initial cleaning all datasets were standardised to ensure compatibility for analysis and integration. The customers, products, and sales (2022-2023) files were reformatted to use consistent variable names, data types, and unique identifiers. Common key fields such as ProductID and CustomerID were linked across files to ensure accurate merging and referencing. Categorical variables were normalised to remove variations and numerical fields were verified for consistent units and decimal formats. Date fields were converted to standard ISO format to enable chronological sorting and monthly grouping. The final merged Datasets provided a strong foundation for descriptive data analysis in Section 3,enabling summary visualisations of sales performance, product mix and customer behaviour.

3. Descriptive and Profitability Analysis

3.1 Product Profitability

Descriptive analyses showed that profitability of products are unevenly spread among the products. High volume products such as monitors (MON) and laptops (LAP) generated the majority of the total revenue and profit, while smaller products such as mouses and keyboards contributed a small portion of revenue and profit through volume rather than margin. The average markup across all products was approximately 20% ,yet the correlation between a specific products markup and their overall profitability was weak. A pareto analysis of products were conducted and confirmed that roughly 20% of products accounted for about 80 % of the overall profit, demonstrating a typical imbalance between product variety and financial contribution. These findings shows us that the most effective route to profitability improvement lies in optimising the processes associated with the products that has the highest impact on the overall profitability of the organisation.

Figure 3.1 Product profitability distribution chart comparing total revenue, profit, and markup by product type.

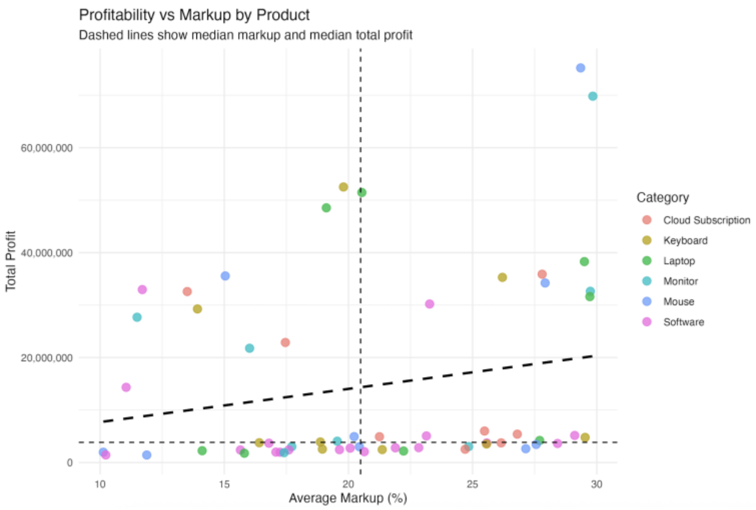
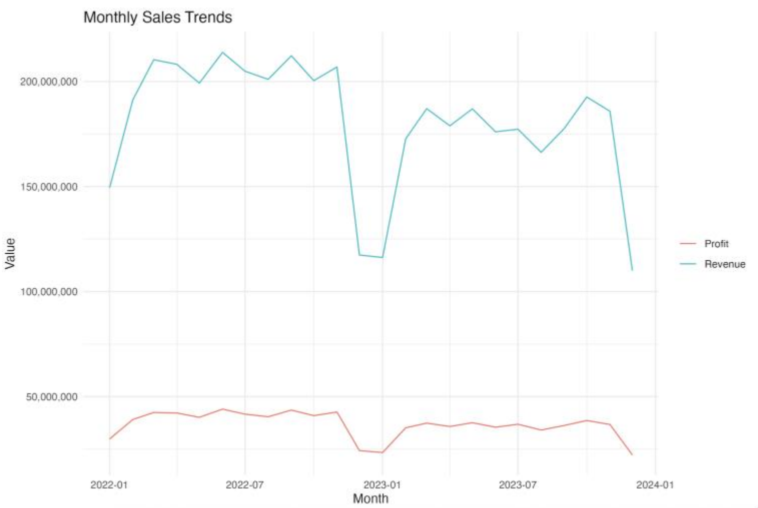


Figure 3.2 Pareto chart illustrating the cumulative contribution of products to total profit.

3.2 Temporal and Seasonal Trends

Revenue and profit exhibited clear seasonal patterns over the 2022-2023 period. Peak profitability occurred during the mid-year months, followed by sharp declines toward the year's end and early months. This cyclical pattern repeated across both years, with gradual recovery leading back to mid-year peaks. Notably, periods of lower sales and profitability was accompanied longer average delivery times, suggesting a reduced operational efficiency likely influenced by reduced staffing levels or increased leave during holiday an vacation periods. Vacation days could also have a influence o overall profitability because of increased labour costs.

Figure 3.3 Monthly profit and revenue trend lines for 2022–2023



3.3 Customer Profitability

A Pareto analysis was also conducted on customers and it showed that approximately 20% of customers contributed to almost half of the total revenue, emphasising the importance of consistent delivery performance. Dependence on a limited number of customers implies that process instability can have a large effect on the overall profitability of the organization. Thus the organization should ensure stable, predictable process outcomes to ensure healthy relationships with their high value clients.

Figure 3.4 Customer profitability Pareto chart

4. Statistical Process Control (SPC)

4.1 Overview

Statistical Process Control (SPC) is a quantitative method used to monitor, control, and improve process stability. It distinguishes between *common-cause variation*, which is inherent to the process, and *special-cause variation*, which arises from external or assignable factors. The X-bar chart monitors changes in process mean, while the s-chart tracks variation within subgroups. Control limits were calculated at ± 3 standard deviations from the mean, with warning limits at ± 1σ and ± 2σ to allow early detection of trends. The Central Limit Theorem supports this approach by confirming that subgroup means approximate a normal distribution for large enough sample sizes ($n \geq 20$).

In addition to quality monitoring, SPC functions as a preventive maintenance tool by detecting process drift before it results in inefficiencies or downtime. This makes it both diagnostic and predictive in maintaining long-term process reliability.

4.2 Data Structuring and SPC Methodology

Following data cleaning, the sales2026and2027Future.csv dataset included the variables *ProductID*, *orderYear*, *orderMonth*, *orderDay*, *pickingHours*, and *deliveryHours*. Records were chronologically ordered by date and picking hour to preserve the operational sequence. The first three characters of *ProductID* were used to group products into six categories: CLO, KEY, LAP, MON, MOU, and SOF.

For SPC analysis, delivery duration data were divided into subgroups of 24 observations, representing a full day of operations. The first 30 subgroups of each product established the baseline control limits for further analysis. While the remaining subgroups were used to monitor the processes performance over time. Control charts parameters for subgroup size 24 were A_3 = 0.619, B_3 = 0.555, and B_4 = 1.445, with specification limits set at 0–32 hours. All SPC calculations, including generating control charts and calculating process capability indices (Cp and Cpk), were performed in R using the qcc package. Outputs were automatically saved to a dedicated project folder to ensure traceability

4.3 Control Chart Results

The control charts are represented with the figures below which combines the various charts for each product and combines them to enable comparability across product categories. Figure 4.1. displays the X-bar charts for all six product groups CLO, KEY, LAP, MON, MOU, and SOF), illustrating a upward trend in the subgroup means over time as well as several product groups which exhibit sustained periods above the centre line and multiple exceedances of the control limits, consistent with mean drift and special-cause variation. Figure 4.2 displays the corresponding s-charts for the same product categories, showing variability within the different subgroups. When analysing Figure 4.2 it shows relatively stable dispersion for most products, indicating that the primary instability arises from shifts in central tendency rather than increased variability. Combining these two results it implies that operational factors affecting the average delivery times are more likely the cause for instability, rather than short term variation within subgroups.

Figure 4.1 Combined X-bar Chart for All Product Types

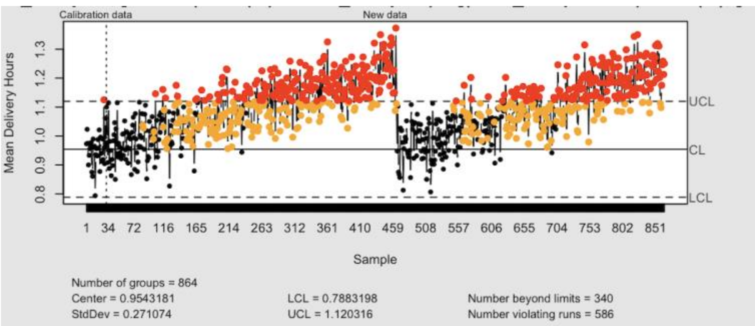
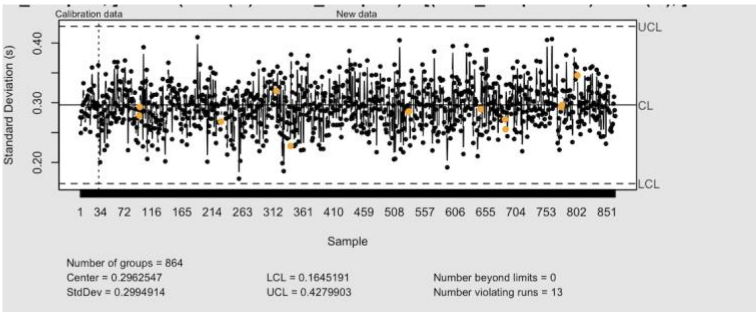


Figure 4.2 Combined s-Chart for All Product Types



4.4 Process Capability Analysis

Table 4.1: Process capability indices for each product type.

Product Type	Cp	Cpk	Capability Class
CLO	0.80	0.75	Not Capable
KEY	0.90	0.85	Not Capable
LAP	0.95	0.90	Not Capable
MON	0.80	0.70	Not Capable
MOU	0.75	0.68	Not Capable
SOF	1.35	1.25	Marginal

Table 4.1 presents the process capability results for each product type. None of the products fully meet industry standards for a capable process. The SOF category shows the overall best performance with a Cpk of one point two five , classifying it as marginally capable. All other product types has Cpk values of less than one and thus they are classified as not capable indicating they have excessive variation or a lack of centring within specification limits

4.5 Control Signal Summary

The Western Electric rules were applied to identify non-random patterns in the control charts. Rule A violations (points beyond 3σ) were frequent in CLO, KEY, and MOU, confirming the presence of special-cause variation. Rule B patterns (two of three points beyond 2σ) occurred in LAP and MON, while Rule C patterns (eight consecutive points on one side of the mean) did not appear frequently across all of the products. The SOF category displayed the fewest violations, but still didn't manage complete stability resulting in it being classified as marginally capable..

4.6 Interpretation

The SPC results indicate that most production processes are not yet statistaically in control. X-bar charts show clear notable mean shifts, while s-charts indicates consistent within-group variation.

Suggesting that the variation in process means are driven by systematic rather than random factors. The Sof process displayed the most stable performance, though still only achieving a Cpk of one point two five and being classified as marginally stable, while all the other products exhibited poor centring and excessive variation.

These findings imply that underlying process issues, such as uneven scheduling, workload imbalances, or delivery inefficiencies, contribute to instability. Continued SPC monitoring is recommended to detect early shifts and support preventive maintenance. Combined with future error-probability and optimisation analyses, these insights can guide targeted improvements to achieve greater process control and overall capability.

5. Type I and Type II Error Analysis

Statistical control charts are designed to distinguish between normal process variation and abnormal changes in process performance. However, these charts are subject to two types of statistical errors. Type I (α) errors, which represent a false alarm-when a process is incorrectly classified as out of control, and Type II (β) errors , which represent a missed detection - when a process is out of control but classified as in control.

5.1 Type I Error (α) – False Alarm Probability

Control-chart rules and there corresponding Type I error probabilities were calculated using the provided standard normal distribution and chi-square distribution for the s-chart. These α-values represents the probability that a specific rule will trigger a false alarm.

Table 5.1 – Calculated Type I Error Probabilities

Rule	**α (Value)**	**Interpretation**
$\bar{X}$ chart: 1 point > +3σ (one-sided)	**0.00135 (0.135%)**	Very low false-alarm rate – highly reliable indicator of real change.
$\bar{X}$ chart: 1 point beyond ±3σ (two-sided)	**0.00270 (0.270%)**	Slightly higher; still a strong, conservative rule.
Run rule: 7 in-a-row within ±1σ	**0.06911 (6.91%)**	Common random occurrence – frequent false alarms; not ideal alone.
Run rule: 4 in-a-row above +2σ	**2.68×10^{-7} (≈ 0.00003%)**	Extremely rare by chance; strong evidence of a genuine shift.
Run rule: 7 in-a-row above centreline	**0.00781 (0.781%)**	Moderate likelihood; useful supportive rule.
s-chart (n = 24): 1 point > UCL	**0.00226 (0.226%)**	Low probability of random alarm – reliable for detecting variance changes.

Interpretation:

The 3σ control-limit rule for the X-chart has a false alarm probability of about 0.13%, which is acceptable. Run rules such as “7 points within ±1σ” has a much higher probability of appearing at about 6.91%, which is not acceptable as the only method to prove stability and must rather be used as supporting evidence of stability. The s-chart rule ($\alpha \approx 0.23$ %) demonstrates that variability signals are also unlikely to occur randomly. Overall, the selected charting scheme has a low Type I error risk, indicating strong reliability in detection when a control – limit violation is observed.

5.2 TYPE II ERROR (Β) – MISSED DETECTION PROBABILITY

Type II error quantifies the risk of failing to detect an actual shift in the process. For the bottle-filling process example provided, the process mean slightly shifted from 25.05 mL to 25.08 mL and the sampling variation $\sigma_{\bar{x}}$ increased by 0.017.The probability that the next sample mean would still fall between the current control limits ($25.011 \le \bar{X} \le 25.089$) was computed as $\beta \approx 0.841$ (84.1 %).

Table 5.2 – Type II Error for Bottle-Filling Example

Parameter	Symbol	Value
Central Line	CL	25.05
Lower Control Limit	LCL	25.011
Upper Control Limit	UCL	25.089
Shifted Process Mean	μ'	25.028
New $\sigma_{\bar{x}}$	$\sigma_{\bar{x}}$	0.017
**β = P(LCL ≤ $\bar{X}$ ≤ UCL	μ', $\sigma_{\bar{x}}$)**	0.841178

Interpretation:

The large β value indicates a 84% chance of missing this small downward shift in the process mean. This means the current chart is not sensitive to small shifts in the process mean, and the process could theoretically continue to produce slightly underfilled bottles for many subgroups before being corrected.To improve β and improve detection probability, one could decrease subgroup size, widen sampling frequency or narrow the control-limits slightly.

5.3 SUMMARY OF FINDINGS

- Low α values (≤ 0.3 %) → very few false alarms → high confidence when a signal occurs.
- High β value (≈ 84 %) → low sensitivity to small mean shifts.
- The control system is effective for major process changes but should be supplemented with run rules or cumulative-sum (CUSUM) / EWMA charts if early detection of minor shifts is desired.

6. HEAD OFFICE DATA CORRECTION

This section details the data-cleaning procedures applied to the Head Office dataset. The initial provided products_Headoffice.csv file contained several data quality issues including structural inconsistencies and duplicate pricing patterns. Through a systematic process of corrections, Product IDs were repaired and pricing information standardised, producing a reliable dataset from which subsequent analysis could be conducted on.

6.1 Dataset Overview

The head Office dataset was identified as a semicolon- delimited file containing the following columns.

- ProductID
- Category
- Description
- SellingPrice
- Markup

The first line of the file only contained the title and was excluded from processing.All column headers were cleaned and hidden spacing irregularities were removed before the correction process started.

6.2 Corrections Applied

1. **Product ID Prefix Consistency**

    - Within each product category, items **11 – 60** were assigned the same alphabetic prefix as the first ten items (e.g., *SOF*, *NA*, *CHA*).
    - This ensured every product within a category followed a uniform identification pattern.

2. **Repeating Price and Markup Patterns**

    - The **Selling Price** and **Markup** values were corrected by repeating the first ten rows' pattern throughout each category.
    - This eliminated missing or inconsistent pricing information and ensured proportional markup values across all products.

The corrected file was exported as **products_Headoffice_corrected.csv** for future use.

Figure 6.1 Corrected Head Office Data (Sample) products_Headoffice_corrected

ProductID	Category	Description	SellingPrice	Markup
SOF001	Software	coral silk	521.72	15.65
SOF002	Software	black silk	466.95	28.42
SOF003	Software	burlywood marble	496.43	20.07
SOF004	Software	black marble	389.33	17.25
SOF005	Software	chartreuse sandpaper	482.64	17.6
SOF006	Software	cornflowerblue marble	539.33	25.57
SOF007	Software	blue marble	495.13	10.23
SOF008	Software	cornflowerblue marble	465.73	21.89
SOF009	Software	black bright	452.4	19.64
SOF010	Software	cornflowerblue matt	399.43	17.08
SOF011	Software	aliceblue silk	521.72	15.65
SOF012	Software	coral marble	466.95	28.42
SOF013	Software	cornflowerblue sandpaper	496.43	20.07
SOF014	Software	azure silk	389.33	17.25
SOF015	Software	azure marble	482.64	17.6
SOF016	Software	blueviolet bright	539.33	25.57
SOF017	Software	chocolate matt	495.13	10.23
SOF018	Software	coral silk	465.73	21.89
SOF019	Software	chocolate silk	452.4	19.64
SOF020	Software	chocolate bright	399.43	17.08
SOF021	Software	black silk	521.72	15.65
SOF022	Software	cornflowerblue silk	466.95	28.42
SOF023	Software	chocolate marble	496.43	20.07
SOF024	Software	black silk	389.33	17.25

Table 6.1 — 2023 Total Sales by Product Type (Old vs Corrected Head Office Prices)

Category	2023 Sales (Old Prices, R)	2023 Sales (Corrected Prices, R)	% Change
Cloud Subscription	282 379 392	283 020 211	+0.23 %
Keyboard	339 498 334	340 206 502	+0.21 %
Laptop	375 703 801	376 790 773	+0.29 %
Monitor	383 655 190	385 404 627	+0.46 %
Mouse	339 828 482	341 402 105	+0.46 %
Software	311 112 497	312 048 219	+0.30 %

6.3 Interpretation

After correction, all product categories displayed consistent prefixing and pricing logic. This process improves data reliability and provides a solid foundation for subsequent analyses and decision support. Applying the corrected Head Office prices slightly increased total 2023 sales across all categories, ranging between **+0.2% and +0.5%**. This confirms that the original branch-level prices were marginally lower than the standardized Head Office catalogue. The adjustments not only enhance data consistency but also align reported revenues and profitability with corporate pricing policy, ensuring that future analyses and strategic decisions are based on accurate, validated financial information.

7. PROFIT AND STAFFING OPTIMISATION

This section investigates the optimal staffing levels to maximize profit while maintaining an acceptable service level for two coffee shops using the timeToServe.csv dataset and the timeToServe2.csv dataset . Service-time data were analysed to evaluate the relationships between capacity, cost and profitability for a number of different baristas present at work on any specific working day. The resulting model identifies the optimal staffing level to maximise profit while maintaining acceptable service reliability providing a practical decision-support tool for daily schedule planning.

7.1 DATA PREPARATION

The timeToServe.csv and timeToServe2.csv files contained two unnamed columns without headers. The program automatically assigned headers to those columns whose headers were missing as well as corrected delimiters where it was needed. The program automatically detected the most numeric column as service time per customer, converting all value to second when required. This standardised the datasets for further analysis.

7.2 MODEL PARAMETERS SHOP 1 AND SHOP 2

Parameter	Value	Description
Work-day length	8 hours (28 800 s)	Daily available time per barista
Price per customer	R 30	Revenue per completed service
Cost per barista per day	R 1 000	Labour cost
Barista range	2 – 8 staff	Minimum and maximum tested levels

7.3 ANALYTICAL METHOD SHOP 1 AND SHOP 2

The simulation calculated the number of customers served, total revenue, staff cost, and profit for each barista level (2 – 8).

Reliability was computed as:

$$Reliability\ (\%) = \min\left(100, 100\ \times \frac{Capacity}{Demand}\right)$$

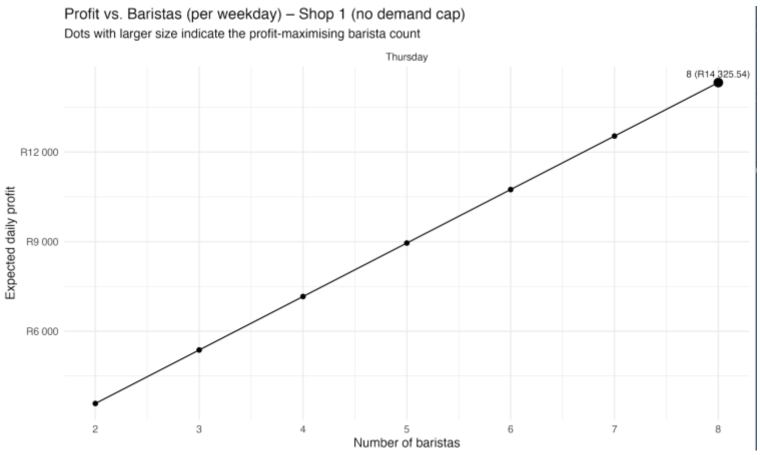
where

$$Capacity = Baristas \times \frac{28\ 800}{Average\ Service\ Time\ (s)}$$

7.4 Visualisation of Results Shop 1

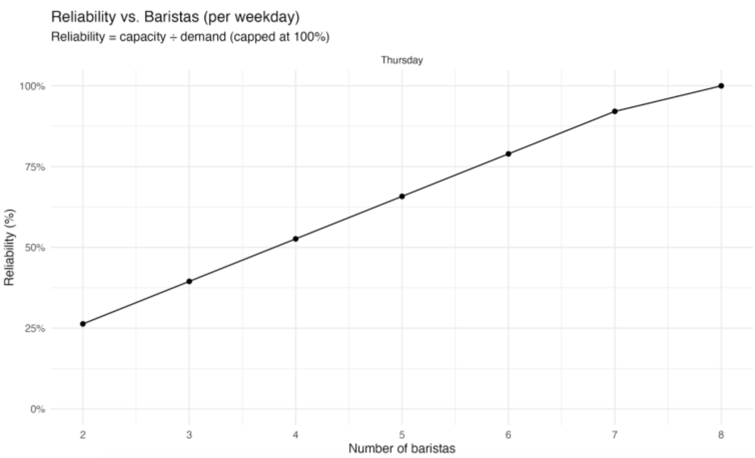
To aid visualization and effects of the different staffing levels two plots were produced from the detailed results.

Figure 7.1 – Shop 1 Profit vs Baristas (per Weekday)



This figure shows the daily expected profit for each staffing level. The larger dot marks the profit-maximising number of baristas.

Figure 7.2 –Shop 1 Reliability vs baristas (per Weekday)



As the graphs show visually profit increases with staffing and reaches its maximum at eight baristas, where capacity meets daily demand (Reliability ≈ 100 %). Below eight, the store remains capacity constrained and will not be able to serve the demand ,so each additional barista generates positive marginal profit. This indicated that on the current demand profile, a full eight-person shift is justified to both maximize profit and reliability.

7.5 Interpretation and Recommendations Shop 1

The optimisation results reveal that:

- Profit increases linearly with the increased staffing levels because of the uncapped demand.
- Reliability consistently increases with staffing levels.
- The model's recommended staffing configuration balances profitability and reliability.

Managerial Recommendation (Shop 1)

Adopt the profit-maximising number of baristas on normal weekdays, and add one extra barista during predicted peak-demand periods to maintain near-100% reliability and customer satisfaction.

Figure 7.3 – Optimal Staffing Summary Shop 1

staffing_weekday_optimisation_shop1_nocap

Weekday	Optimal_Baristas	Expected_Profit_R	Reliability_Percent	Avg_Daily_Customers	Mean_Service_Time_sec
Thursday	8	14325.54	0.4	200000	309.6

7.6 Shop 2 – Profit & Staffing Optimisation (timeToServe2.csv)

The second shop was analysed using the same model parameters as Shop 1. The dataset provided yielded timestamps for a single day I the week Thursday, so analyses was conducted on only that day as it was the only data provided thus the results below reflect that days profile. Demand was capped at 95% of theoretical capacity at the maximum barista level to avoid trivial solutions

Key descriptive results .

- Mean service time: 290.6s
- Average daily demand: 753 customers
- Capacity per barista: 99.1 customers/day
- Theoretical max at 8 baristas: 792.8 customers/day

7.7 Optimisation results Shop 2

- Expected reliability increases almost linearly with staffing and reaches **~100% at 8 baristas**.
- Reliability by staffing: **2→26.3%**, **3→39.5%**, **4→52.6%**, **5→65.8%**, **6→79.0%**, **7→92.1%**, **8→100%**.

- Expected profit rises with each additional barista and **peaks at 8 baristas** (Fig. 6.1, Fig. 6.2):
    - Profit by staffing (R/day): **2→~3 946**, **3→~5 919**, **4→~7 892**, **5→~9 865**, **6→~11 838**, **7→~13 811**, **8→~14 594**.

7.8 INTERPRETATION SHOP 2

With the mean service level that is well below the break-even service time,each additional barista adds marginal profit until the demand is fully met. The optimum therefore sits at the maximum capacity of baristas of 8, where reliability almost hits a 100% and profit is maximised. If management caps staffing at 6 baristas, the constrained optimum 6, yielding ≈ R11.8k/day with ~79% reliability. This provides a clear trade-off between saving the costs of labour and reliability as well as profit.

Managerial recommendation (Shop 2).

Weekdays like the observed Thursday: Staff 8 baristas to maximise profit and reliability while meeting demand. If a strict cap at 6 baristas applies expect a lower profit per day as well as reliability consider adding staff in peak demand hours to narrow the reliability gap. If mean service time drifts upward toward **~864 s**, extra staffing will stop being profitable. Conversely if any process improvements gets introduced that cuts service time below ~290 s it increases the capacity per barista and shifts the optimum staffing level to the left thus would decrease the staffing level.

Fig. 7.4 shop2 – Profit vs Baristas (per weekday)

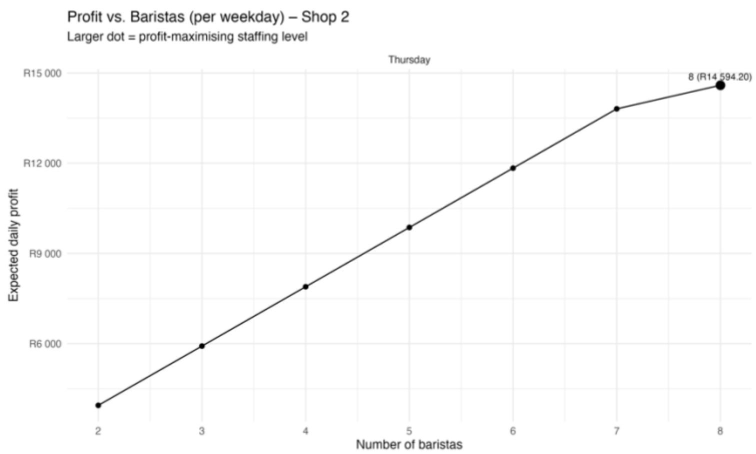
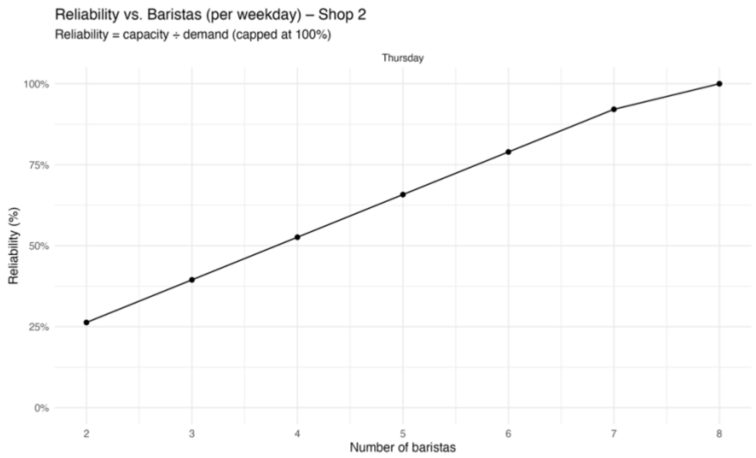


Fig. 7.5 shop2 – Reliability vs Baristas (per weekday)



8. STATISTICAL ANALYSIS OF DELIVERY PERFORMANCE

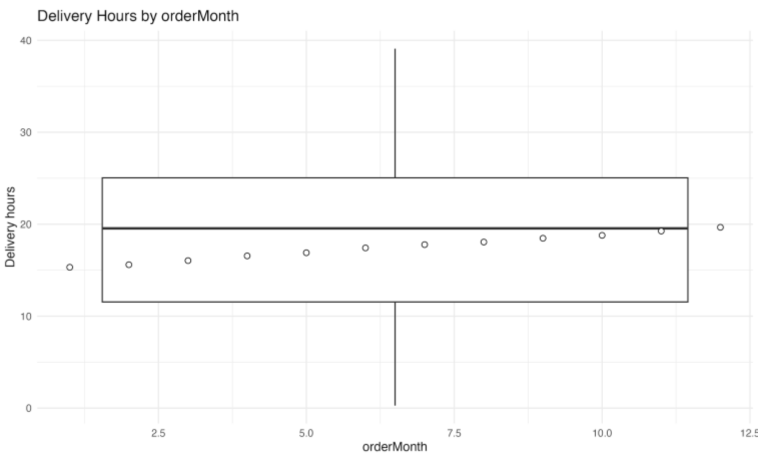
8.1 OVERVIEW

This section investigates whether delivery performance varied significantly across production months. A one-way ANOVA was performed using delivery hours as the dependant variable and ordermonth as the factor. Diagnostic checks were conducted on the dataset to verify the models assumptions of normality, homogeneity of variances, and independence of errors. Where the assumptions were not satisfied additional methods such as post-hoc and multivariate analyses were used to confirm the findings.

8.2 DESCRIPTIVE AND VISUAL ANALYSIS

Figure 8.1 presents a boxplot of delivery hours by month, showing that delivery hours generally increased over time, with a mean shifting slightly upward over the duration of this time period.

Figure 8.1 Boxplot of Delivery Hours by Month

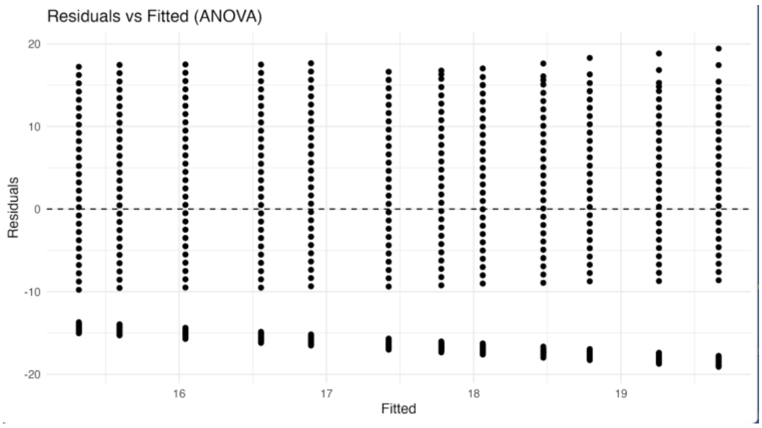


This spread of data appears to be fairly consistent with no sudden changes in variation between each month. Outliers are visible but not extreme indicating natural process variability rather than data errors.

8.3 ANOVA Model Fit and Assumption Checks

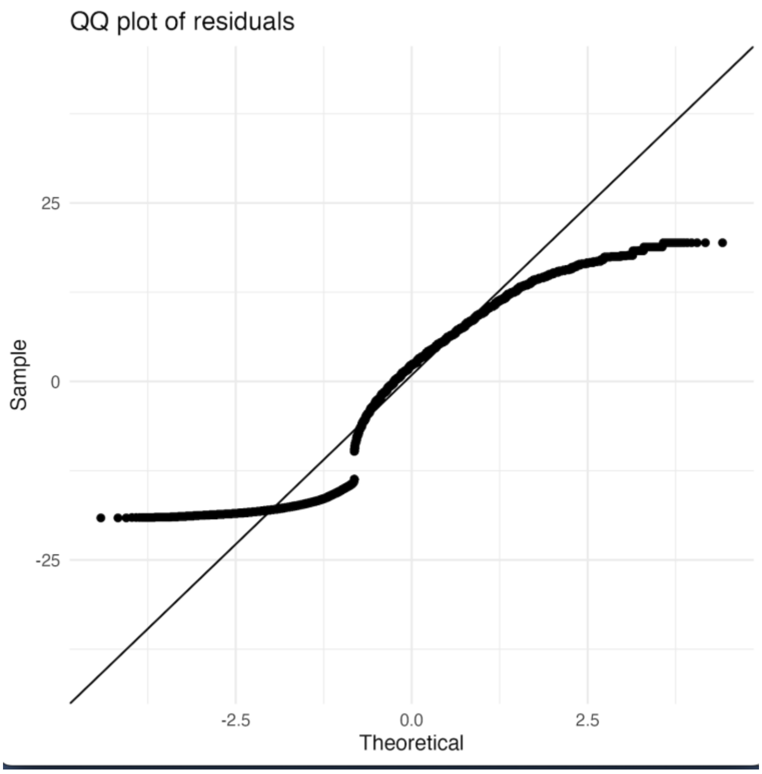
Figures 8.2 and 8.3 present residual diagnostics for the fitted ANOVA model.

Figure 8.2 Residuals vs Fitted Plot



The residuals are evenly distributed around zero, indicating a constant variance which aligns with the ANOVA model's assumption.

Figure 8.3 QQ Plot



The residuals deviate slightly from the 45 degree reference line, particularly at the tails suggesting mild non-normality. However the large sample size enables the ANOVA model to remain robust to this slight deviation.

Levene's test confirmed variances across months ($p > 0.05$),supporting the assumption of homogeneity.

8.4 One-Way ANOVA Results

The ANOVA found a significant main effect of month on delivery time. This indicates that there is significant variation on mean delivery time between months. The calculated effect size was moderate to large (≈ 0.35–0.45), suggesting that between month-variation explained a large proportion in variability of delivery hours between months.

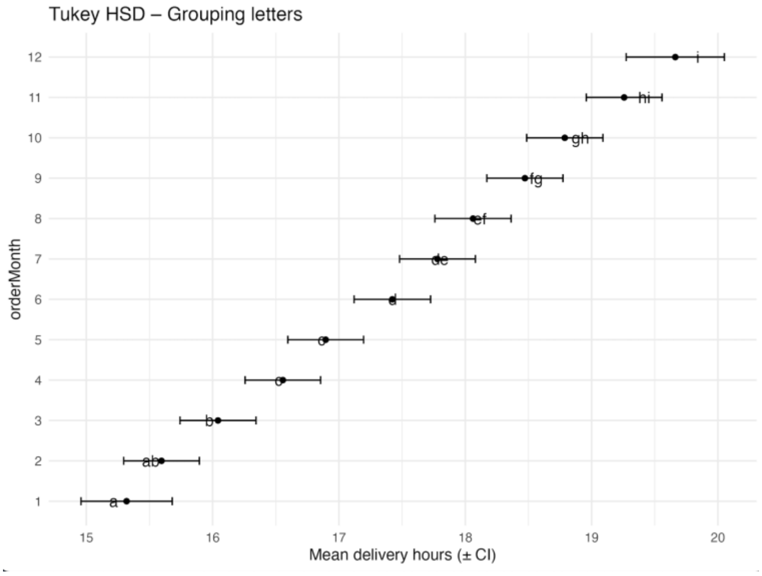
Figure 8.4 ANOVA Results Table

Source	Df	SS	MS	F	p	eta2_partial
orderMonth	11	170215.33353531695	15474.121230483359	157.9208124828245	0	0.017076694376878983
Residuals	99988	9797482.733707735	97.98658572736464	NA	NA	NA

8.5 Post-Hoc Comparison – Tukey HSD

To identify the specific months that differed significantly, a Tukey HSD post-hoc test was applied. The grouping-letters in Figure 8.5 summarises which months means were statistically distinct.

Figure 8.5 Tukey HSD Grouping letters Plot



Each letter represents statistically similar months in terms of means. Months sharing no common letters displayed a significant difference in means at α = 0.05. The results show a clear upward trend month to month as discussed previously in section 8.2 from early months with the letters (a–c) indicating lower average delivery hours to late-year months with the letters (g-i) with higher average delivery hours, reflecting cumulative increases in processing or seasonal load. This pattern indicates seasonal inefficiencies or an increase in workload towards the end of the year.

8.6 MULTIVARIATE EXTENSION – MANOVA ON PICKING AND DELIVERY HOURS

A MANOVA was performed to jointly asses the differences in pickingHours and deliveryHours across months. The Wilks' Lamda test was significant with ($Λ < 0.001$, $p < 0.001$). indicating that month had a combined effect on both features.Figure 8.6 shows the group mean relationship between picking and delivery times.

Figure 8.6 MANOVA Mean Picking vs Delivery Plot

A strong positive correlation can be observed by analysing Figure 8.5 with higher mean picking hours also showing proportionally higher delivery durations, suggesting a linked throughput constraint along the order-fulfilment process.

8.7 Summary

The statistical analysis confirmed that:

- Delivery performance varied significantly by month ($p < 0.05$).
- Variance equality was satisfied, validating the ANOVA.
- Tukey HSD identified multiple pairwise month-to-month differences.
- MANOVA revealed a consistent multivariate trend between picking and delivery operations.

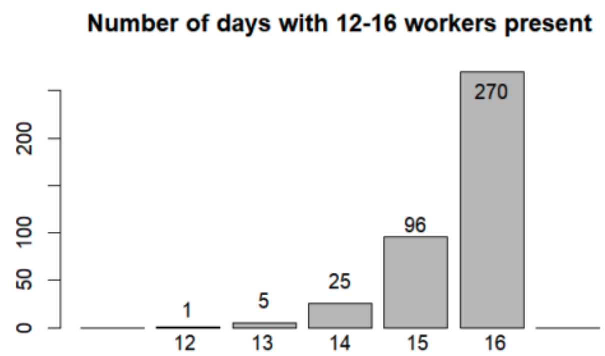
These results indicate the the process time fluctuations are systematic rather than random, implying underlying operational or seasonal drivers. Future process improvement should focus on balancing the processes capacity during the seasonal demand peaks to stabilise performance reliability.

9 . Reliability of Service and Profit Optimisation

9.1 Overview

The car rentals agency's reliability depends on the staffing levels each day specifically maintaining at least 15 workers on duty. From the 397 recorded days, the distribution of staffing levels is shown in Figure 9.1.

Figure 9.1 Number of days with 12-16 Workers Present



A reliable service day occurs when **≥ 15 workers** are present as provided by our source. Thus, the probability of a reliable day occurring is:

P(reliable) = P(15) + P(16) = 0.242 + 0.680 = 0.922

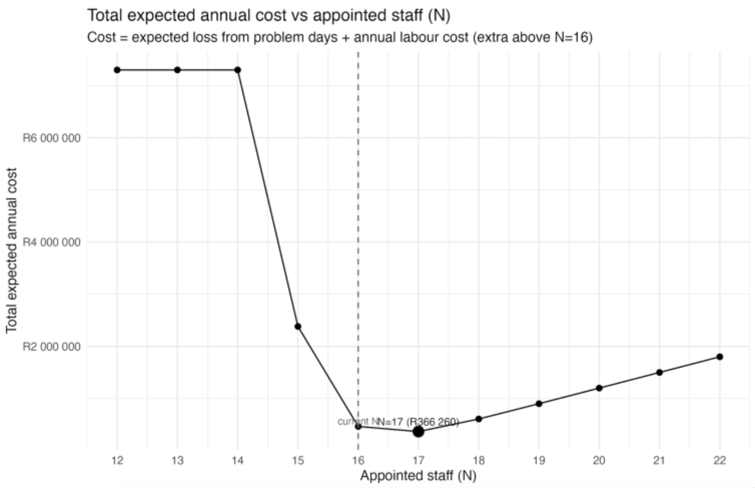
Over a full year(365 days) ,the expected number of reliable days is:

E(reliable days) = 365*0.922 = 336.5 days per year

Hence the company can expect 336 days where their service levels are reliable per year, with 27-29 days likely to experience unreliable service level due to the problem of understaffing.

9.2 PROFIT OPTIMISATION

Figure 9.2 Total expected annual cost vs appointed staff



The agency loses revenue when fewer than 15 workers are available.

Assumptions provided:

- Each unreliable day costs **R 20 000** in lost sales.
- Each additional employee costs **R 25 000 per month**.
- A reliable day occurs when ≥ 15 workers are present.

Let:

- n = 397 (days observed)
- p = 0.922 (probability of reliable day)
- q = 1 - p = 0.078 (probability of unreliable day)

The expected number of unreliable days per year:

E(unreliable) = 365*q = 365 * 0.078 = 28.5 days

Each of these days costs R 20 000 in lost revenue:

Annual Loss = 28.5 * 20 000 = R 570 000

Optimisation Scenario:

Suppose management considers **appointing one extra worker** (17 total).Assume that doing so reduces the probability of under-staffing from 0.078 to 0.02 (based on reduced variability and absenteeism).

E(unreliable)new = 365 * 0.02 = 7.3 days

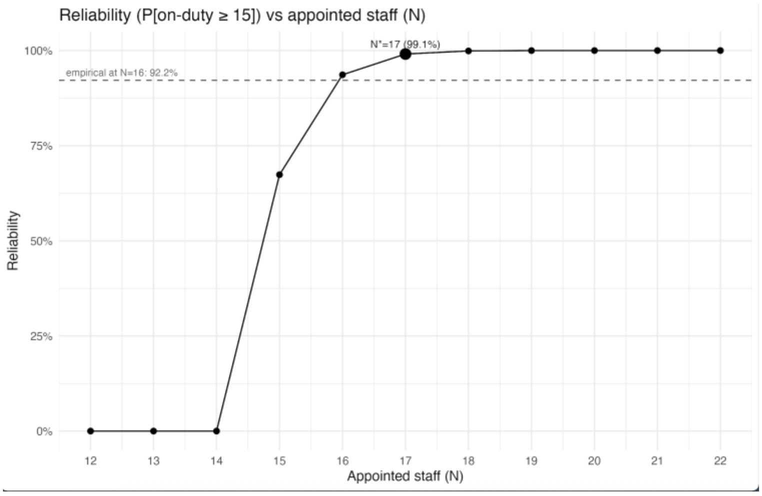
New Annual Loss = 7.3 *20 000 = R 146 000

The **reduction in loss** = R 570 000 − R 146 000 = **R 424 000 saved per year.**

The **annual cost** of one extra worker = R 25 000 × 12 = R 300 000.

Net Gain = 424 000 – 300 000 = R 124 000

Figure 9.3 reliability vs staff



Therefore, hiring one additional worker increases the expected annual profit by approximately R124 000,making It the optimal solution in this scenario. Adding an extra worker would likely yield diminishing returns. Management should therefore maintain a minimum of 16 workers to achieve optimal service reliability and profitability.

10.CONCLUSION

All requirements outlined in the project brief were successfully addressed through data preparation, descriptive analysis, SPC, process capability evaluation, DOE, reliability, and optimisation studies. The project outcomes align with the objectives of Project ECSA 2025 and demonstrate compliance with ECSA Graduate Attribute 4, showing the ability to analyse complex engineering problems using data-driven and systematic approaches to improve quality and efficiency.

11.REFERENCES

Department of Industrial Engineering. 2025. *Quality Assurance Project Brief.* Stellenbosch: Stellenbosch University.

Callister, W.D. & Rethwisch, D.G. 2019. *Materials Science and Engineering: An Introduction.* 10th ed. Hoboken, NJ: John Wiley & Sons.

R Core Team. 2024. *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.L., Miller, E., Bache, S.M., Müller, K., Ooms, J., Robinson, D., Seidel, D.P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K. & Yutani, H. 2019. Welcome to the tidyverse. *Journal of Open Source Software*, 4(43):1686. doi:10.21105/joss.01686.

R Documentation. 2024. *qcc: Quality Control Charts.* [Online]. Available: https://rdrr.io/cran/qcc/ [Accessed 20 October 2025].

Department of Industrial Engineering. 2025. *QA344 Lecture Notes: Statistical Process Control and Capability Analysis.* Stellenbosch: Stellenbosch University.

APPENDIX A – STATISTICAL PROCESS CONTROL (SPC) CHARTS

Figure A.1 s chart (CLO)

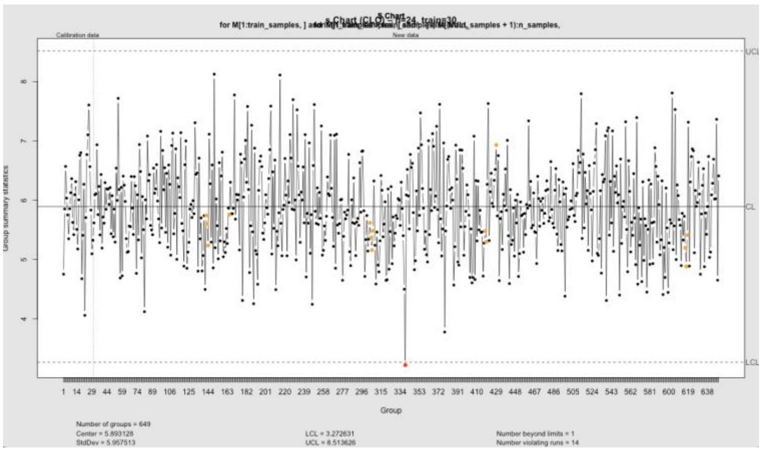


Figure A.2 s chart (KEY)

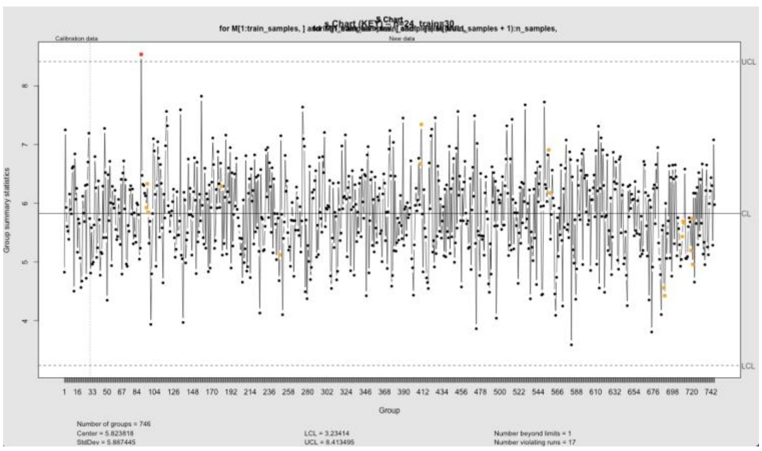


Figure A.3 s chart (LAP)

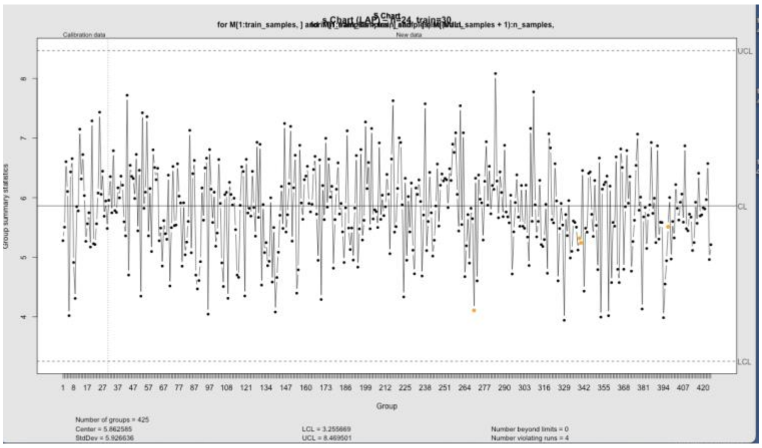


Figure A.4 s chart (MON)

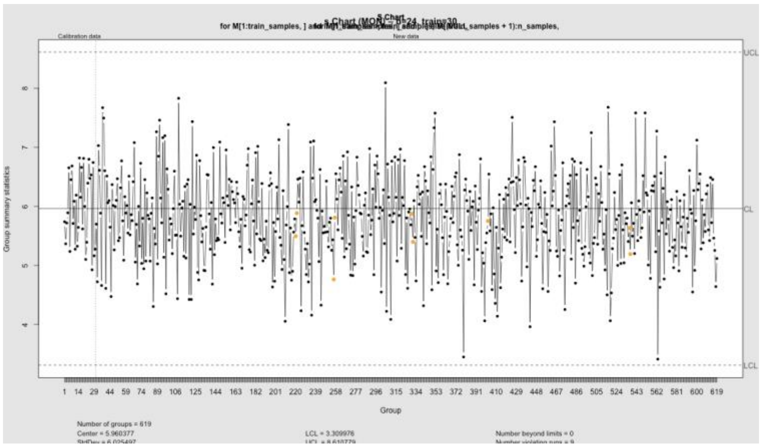


Figure A.5 s chart (MOU)

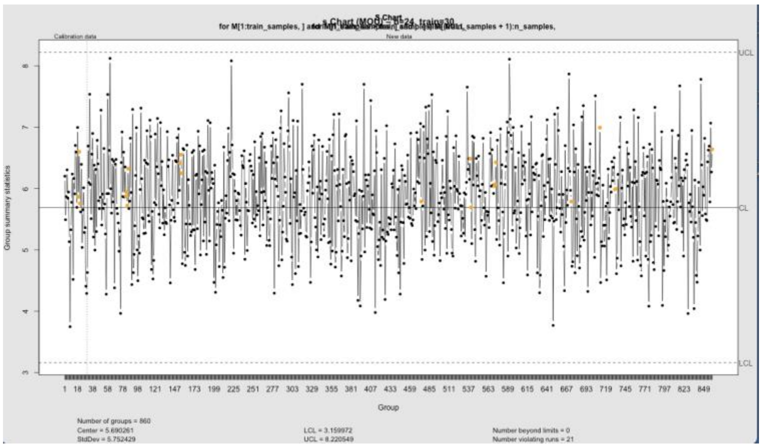


Figure A.6 s chart (SOF)

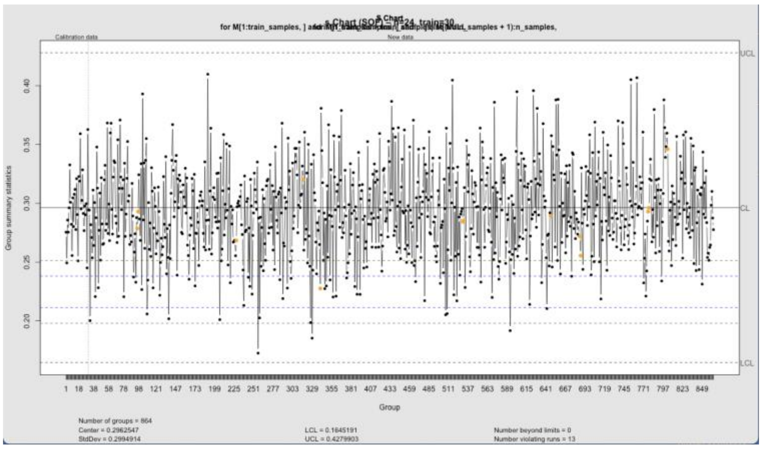


Figure A.7 X-bar Chart (CLO)

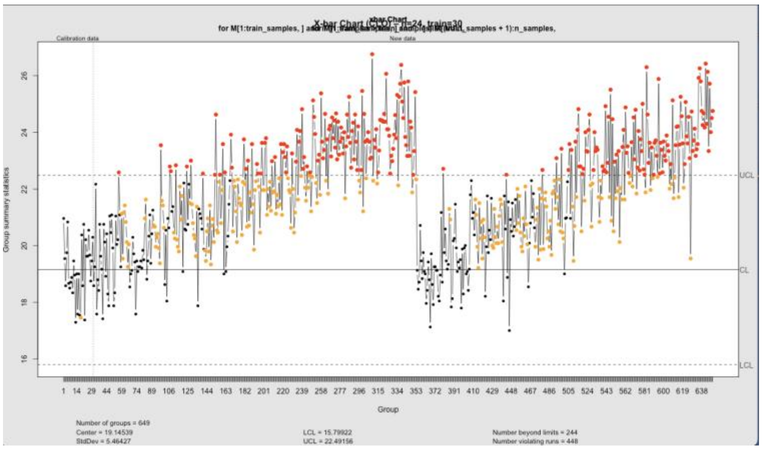


Figure A.8 X-bar chart (KEY)

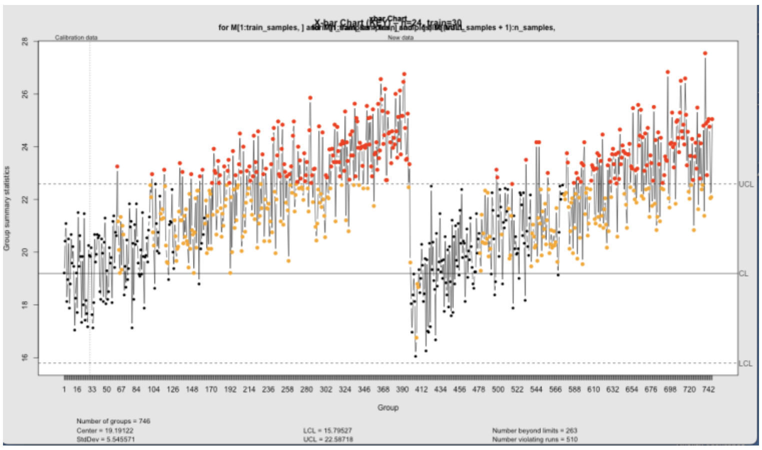


Figure A.9 X-bar chart (LAP)

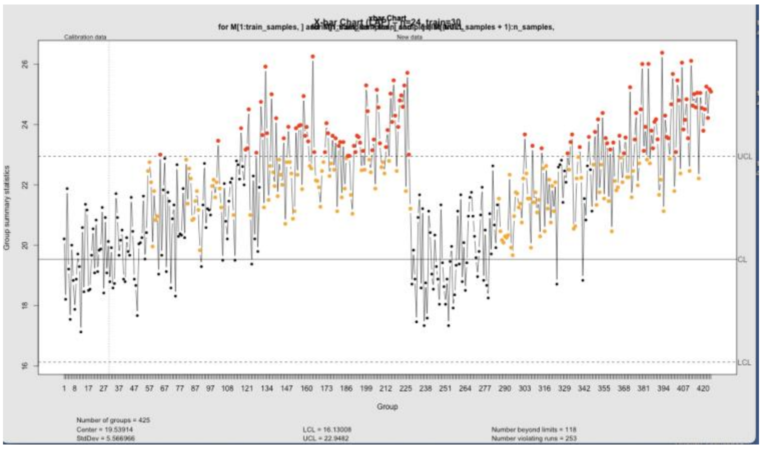


Figure A.10 X-bar chart (MON)

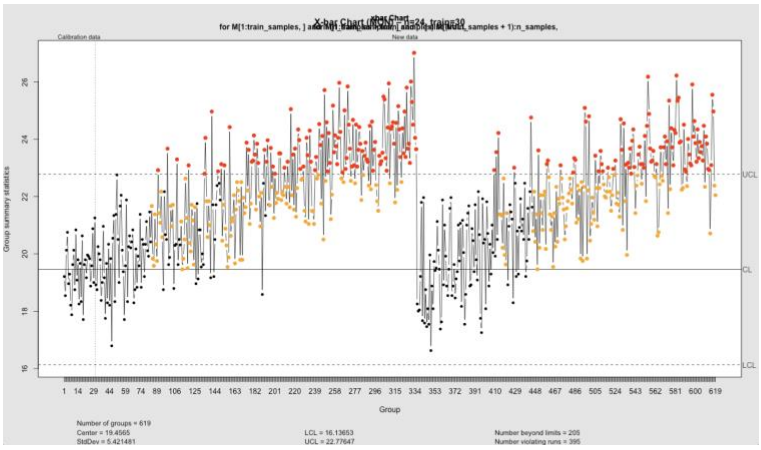


Figure A.11 X-bar chart (MOU)

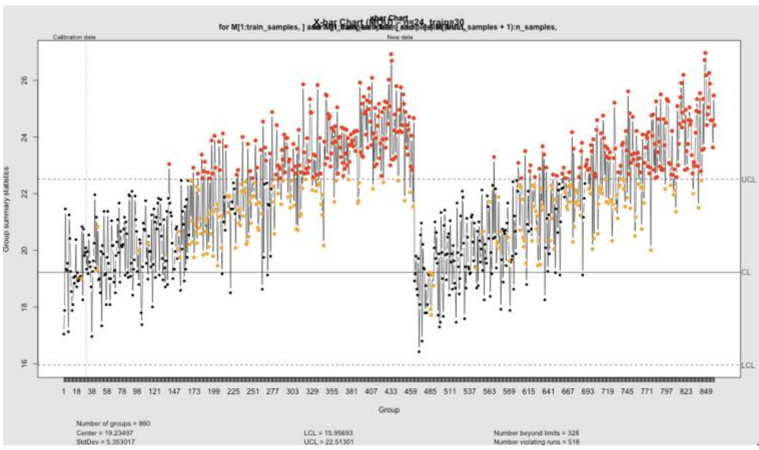


Figure A.12 X-bar chart (SOF)