10/24/2025

ENGINEERING COUNSEL OF SOUTH AFRICA REPORT

Quality Assurance 344

Stellenbosch
UNIVERSITY
IYUNIVESITHI
UNIVERSITEIT

Mila Rossouw - 27018555
STELLENBOSCH UNIVERSITY

Table f Contents

Table of Tables

Table of Figures

Abstract

This report presents a comprehensive quality assurance (QA) investigation conducted in alignment with ECSA GA4 outcomes, integrating statistical analysis, process control and optimization techniques. Using R statistical software, multiple datasets—including products_data2025.csv, customers_data.csv, products_Headoffice.csv, sales20222023.csv, sales202and2027Future.csv, timeToServe.csv and timeToServe2.csv —were analysed to evaluate process performance, customer segmentation and service reliability. The study applies descriptive statistics, control charts, process capability indices, ANOVA and binomial probability modelling to assess operational efficiency and identify improvement opportunities. Results highlight significant corrections to data integrity, revealing shifts in product pricing accuracy and uncovering deficiencies in process capability ($Cpk < 1.0$) across most categories. Additionally, optimization models demonstrated that increasing staffing levels yields measurable cost savings and improved reliability in service delivery contexts. Overall, the report underscores the importance of integrating quantitative quality tools with practical decision-making to achieve sustainable process stability and customer satisfaction.

Introduction

This report applies quality assurance and statistical analysis techniques to evaluate process performance using real industrial datasets. It covers descriptive statistics, Statistical Process Control (SPC), process capability assessment and error estimation to determine product consistency and system stability. Additional analyses include workforce optimization and service reliability evaluations that aim at improving operational efficiency and profitability. The study demonstrates the application of data-driven decision-making consistent with ECSA GA4 outcomes in Industrial Engineering.

Part 1: Descriptive Statistics and Analysis

1.1 Introduction and Methodology

The analysis was performed in R using datasets products_data2025.csv, customers_data.csv and sales_data.csv. These files were imported through the read.csv() and read_excel() functions, with verification of variable types via str() and summary(). The dplyr package was used to clean and manipulate data, removing duplicates, standardizing categorical variables (e.g., "Laptop," "Monitor," "Mouse). Visualizations were created using ggplot2 bar and scatter plots (geom_bar(), geom_point()) to reveal sales patterns and price dispersion across categories. The resulting descriptive analysis provided essential baseline insight into product performance and pricing variability before capability and control analyses.

1.2 Product Portfolio Analysis

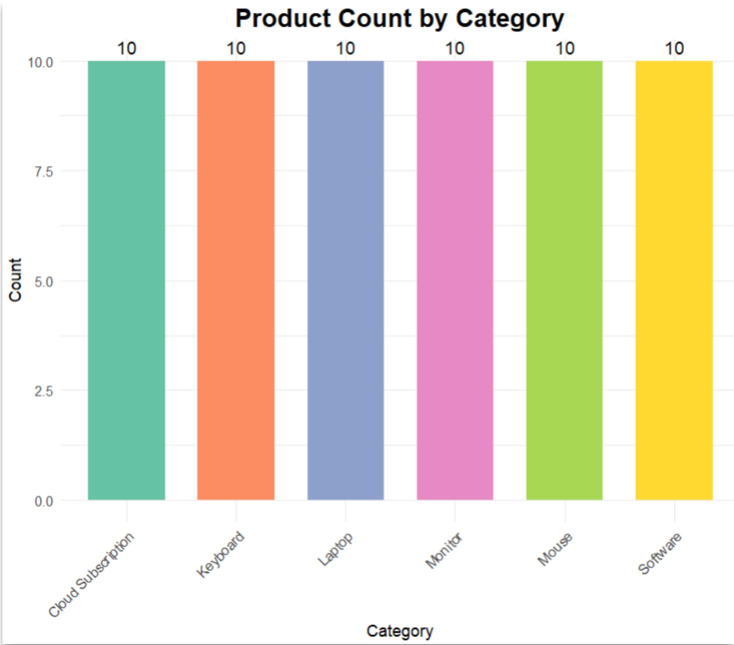

Figure 1: Production count by category

The product dataset includes six categories: Cloud Subscription, Keyboard, Laptop, Monitor, Mouse and Software. Each category contains ten products, totalling sixty items, indicating uniform distribution across categories.

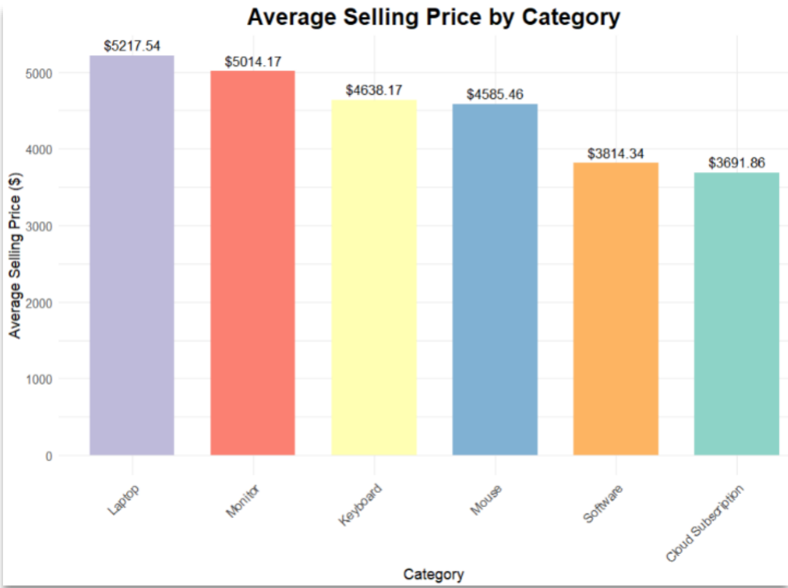

Figure 2: Average Selling Price by category

Average selling prices vary slightly between categories. Laptops have the highest average selling price at $5,217.54, followed by Monitors at $5,014.17. Software and Cloud Subscription have the lowest averages at $3,814.34 and $3,691.86, respectively. These minor variations suggest consistent pricing practices.

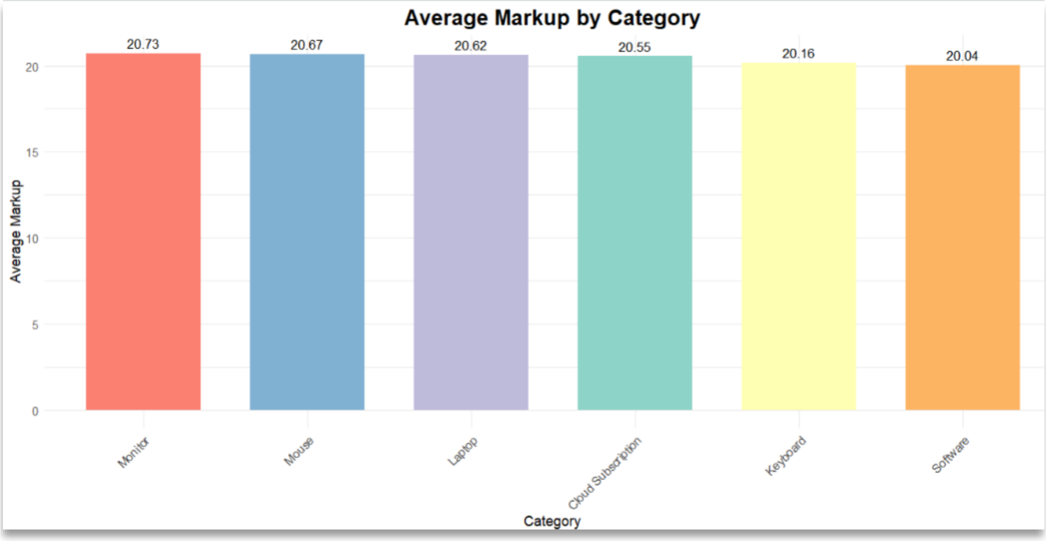

Figure 3: Average Markup by category

Markup percentages range narrowly from 20.04 to 20.73, showing minimal variation in profit margins. This uniformity indicates stable pricing and consistent financial management across all product categories

1.3 Customer Demographic Analysis

Customer data shows uniform distribution across cities, with customer counts ranging from 647 to 780. Average income and demographic patterns appear evenly spread, showing no extreme disparities. The sales data from 2022 and 2023 reveal a slight overall decline in total sales, though trends remain similar year to year.

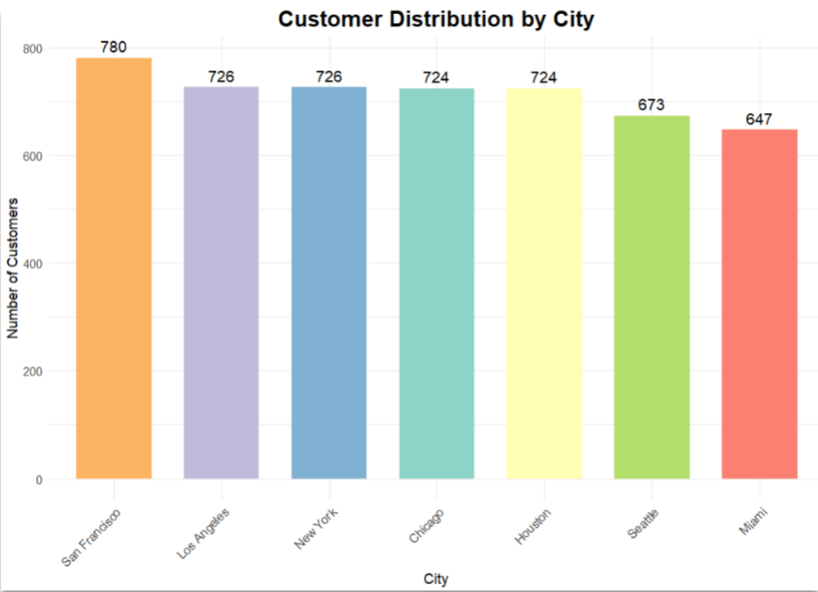

Figure 4: Customer distribution by city

The first chart of customer data illustrates customer distribution across seven cities, with the city names on the x-axis and the number of customers on the y-axis. The range of customers, from a minimum of 647 to a maximum of 780, appears relatively consistent, suggesting a uniform distribution across these cities.

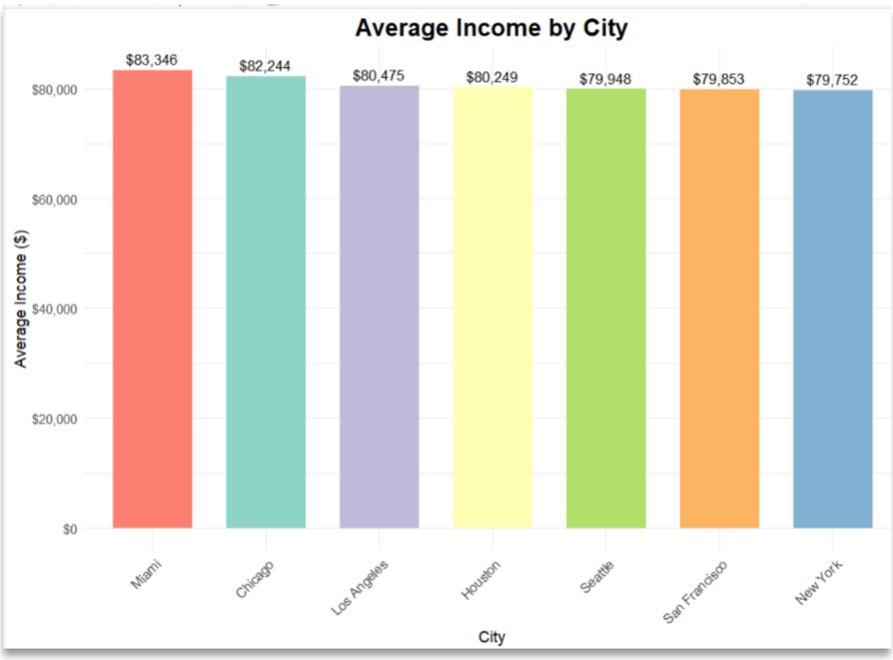

Figure 5: Average Income by City

The second chart displays the average income in dollars on the y-axis against cities on the x-axis. The income levels are evenly spread, with only slight variations and minimal peaks, indicating a uniform distribution across the cities.

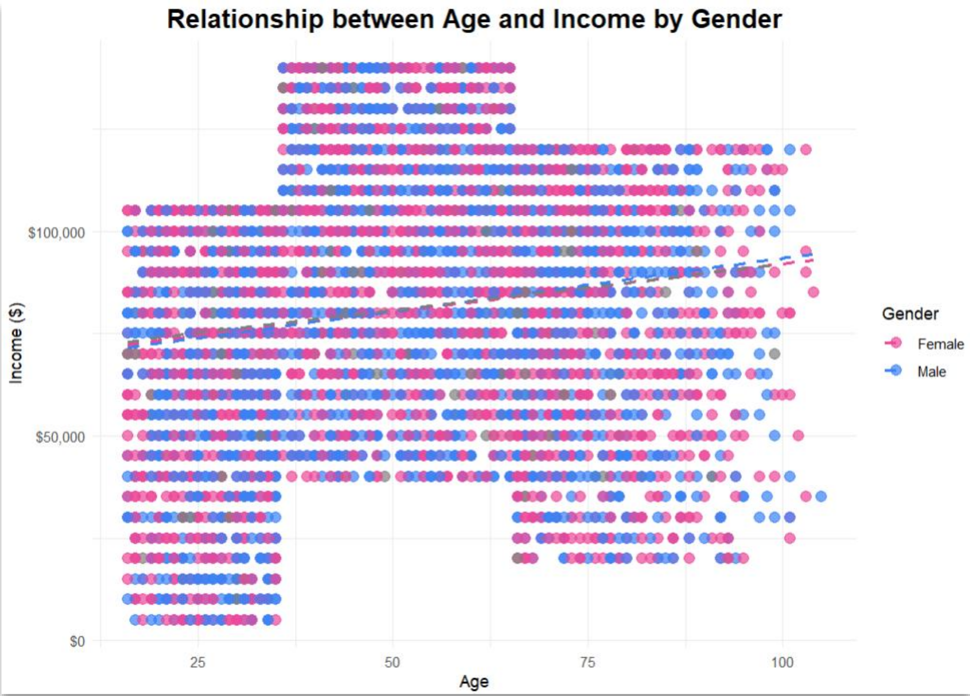

Figure 6: Relationship between Age and Income by Gender

The chart shows that income levels do not consistently rise with age. Both men and women earn similar amounts overall, though slightly more men appear in the higher income range. Younger people tend to have more uniform earnings, while income differences become wider among older age groups.

1.4 Sales Performance and Trend Analysis

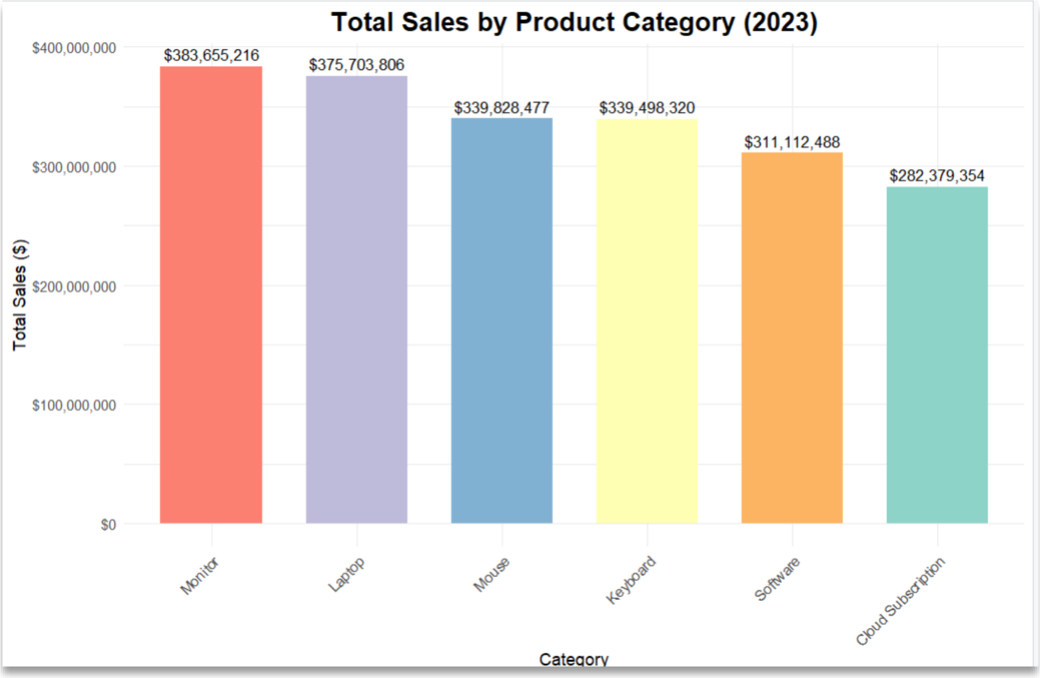

Figure 7: Relationship between Age and Income by Gender

The chart presents total sales in dollars on the y-axis, categorized by product type on the x-axis. The data exhibits a gentle downward trend, though the differences between categories remain relatively small.

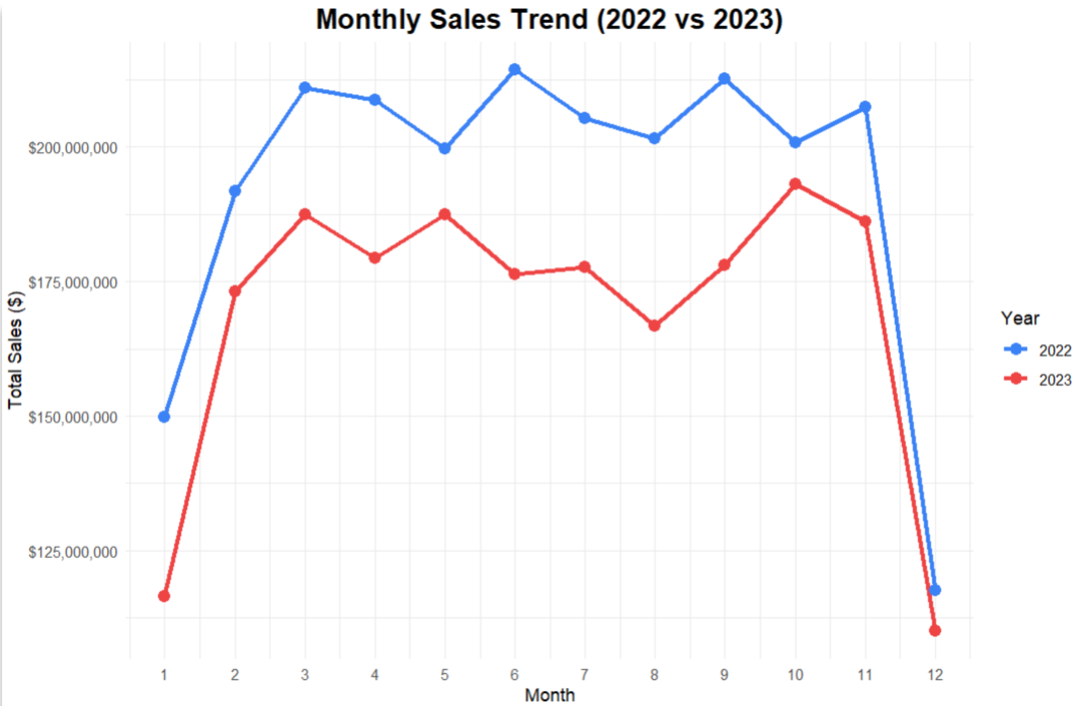

Figure 8: Monthly Sales Trend (2022 vs 2023)

The first chart displays monthly sales trends for 2022 and 2023, with months on the x-axis and total sales on the y-axis, represented by blue lines for 2022 and red lines for 2023. Overall, 2023 shows lower total sales compared to 2022, with a similar trend except for a notable peak in month 10, which was absent the previous year.

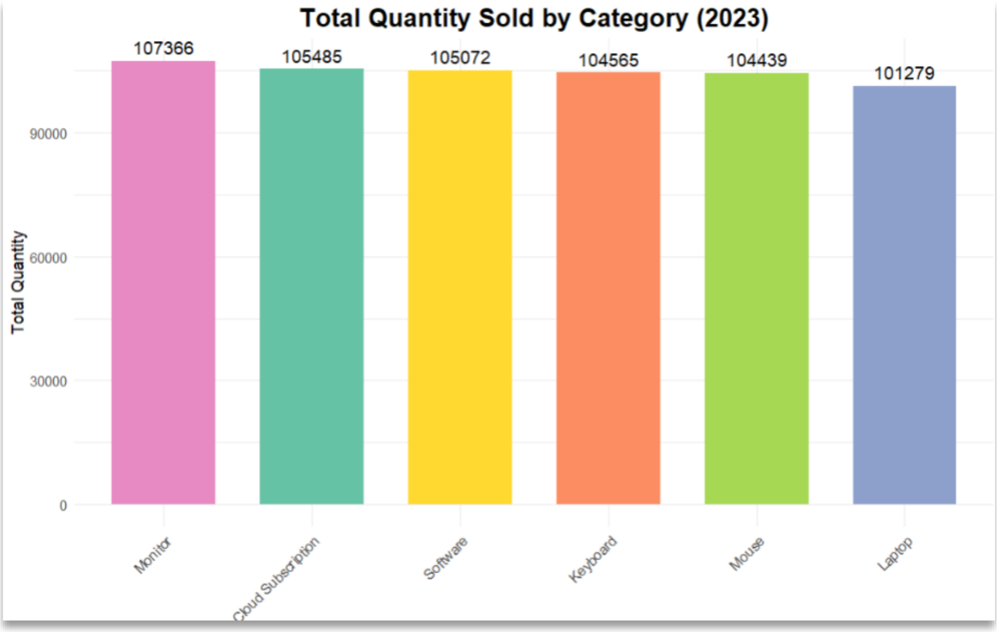

Figure 9: Total quantity sold by category (2023)

The second chart illustrates total quantity sold in 2023 across various categories on the x-axis and quantity on the y-axis. The distribution appears uniform across all categories.

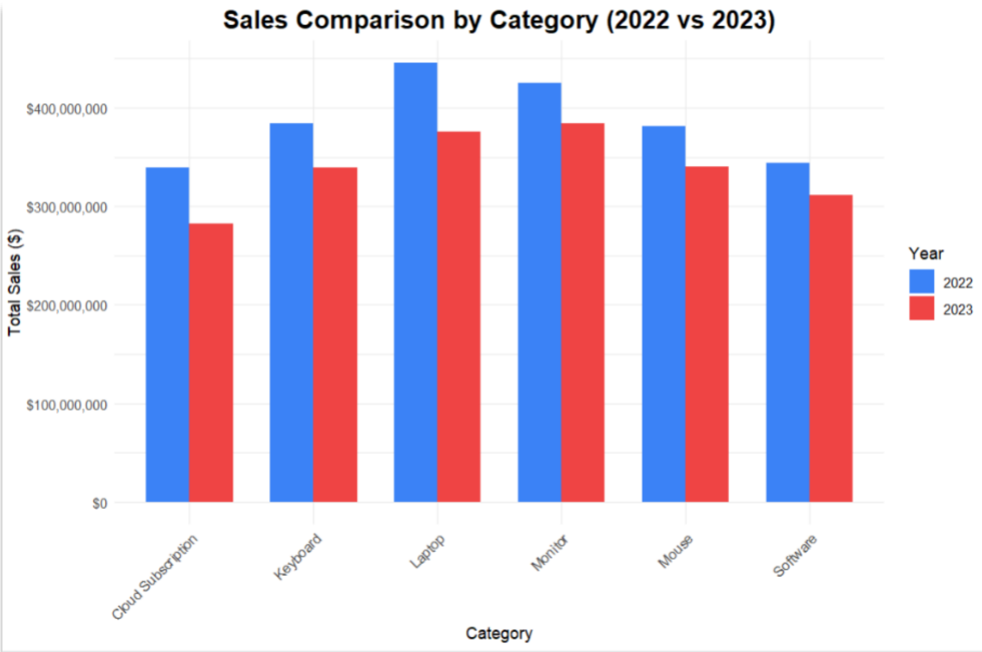

Figure 10: Sales Comparison by Category (2022 vs 2023)

The third chart compares sales for 2022 and 2023, with categories on the x-axis and total sales in dollars on the y-axis, using blue bars for 2022 and red bars for 2023. It indicates lower total sales across all categories in 2023 compared to 2022.

Part 3: Statistical Process Control and Capability

3.1 Methodology:

Process control analysis was carried out using delivery time column in the dataset. Subgroup means ($\bar{X}$) and standard deviations were calculated for each product type using aggregate() and sd() functions. The qcc and ggplot2 packages were used to construct $\bar{X}$- and s-charts, applying constants A_3, B_3, and B_4 consistent with a sample size of 24. Control limits were computed manually in R using $UCL = \bar{X} + A_3 s$ and $LCL = \bar{X} - A_3 s$. The rule detection was incorporated to identify non-random variation patterns like four consecutive points beyond ±2σ. Capability indices $C_p, C_{pk}, C_{pu}, C_{pl}$ were computed relative to the Voice of the Customer limits (LSL = 0, USL = 32). This combination of control charting and capability evaluation provided a quantitative assessment of process stability and customer requirement conformance.

3.2: Control Chart Initialization

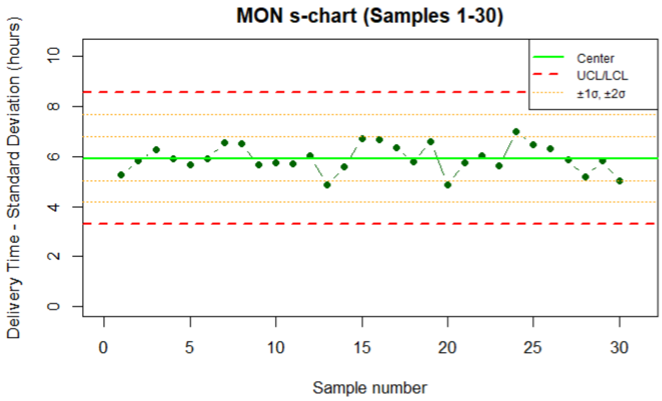

Figure 11: Monitor s-chart for initial 30 samples

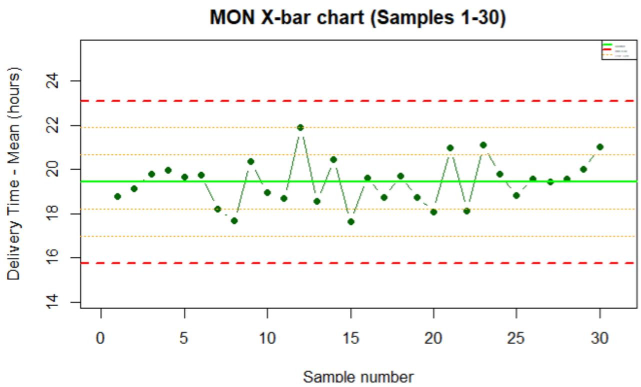

Figure 12: Monitor X-bar chart for initial 30 samples

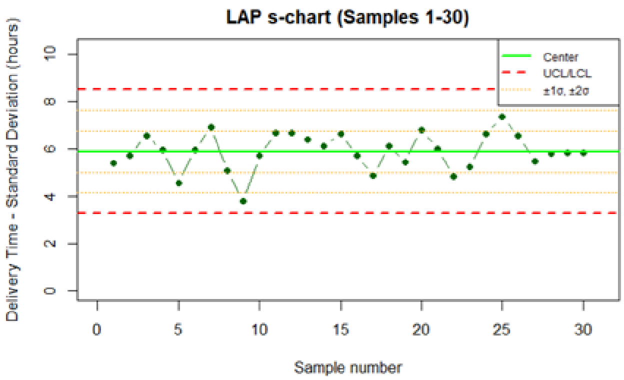

Figure 13: Laptop s-chart for initial 30 samples

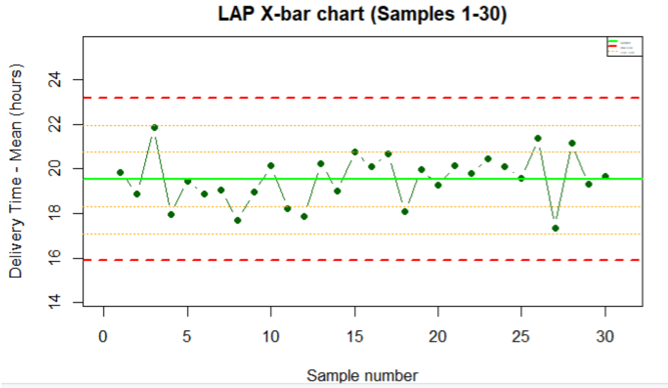

Figure 14: Laptop X- bar chart for initial 30 samples

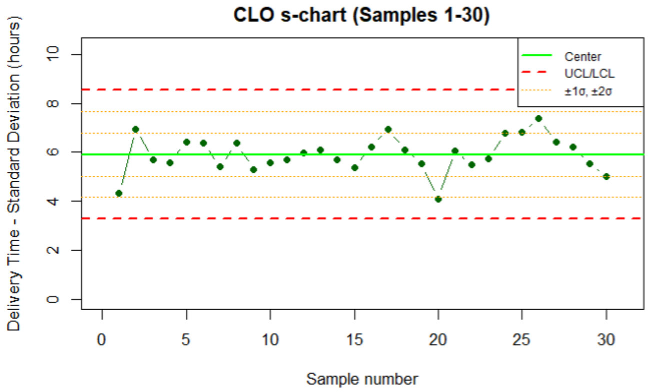

Figure 15: Cloud Subscription s-chart for initial 30 samples

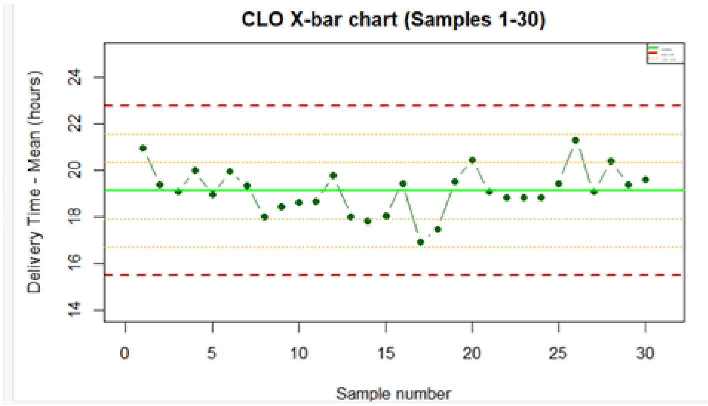

Figure 16: Cloud subscription X-bar chart for initial 30 samples

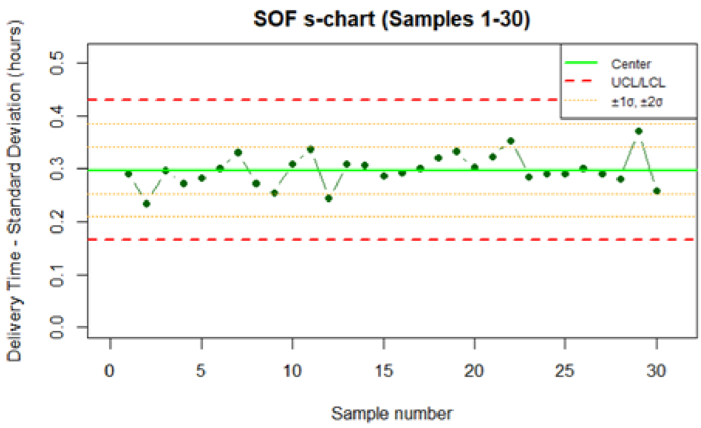

Figure 17: Software s-chart for initial 30 samples

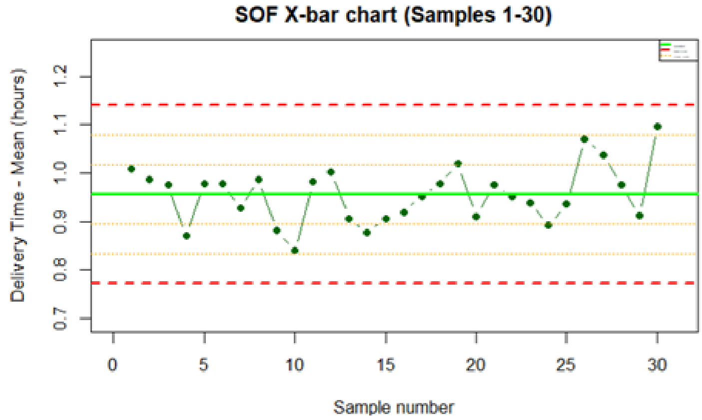

Figure 18: Software X-bar chart for initial 30 samples

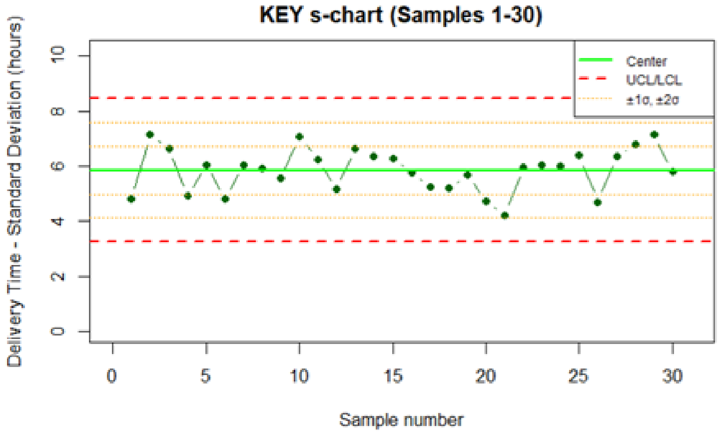

Figure 19: Keyboard s-chart for initial 30 samples

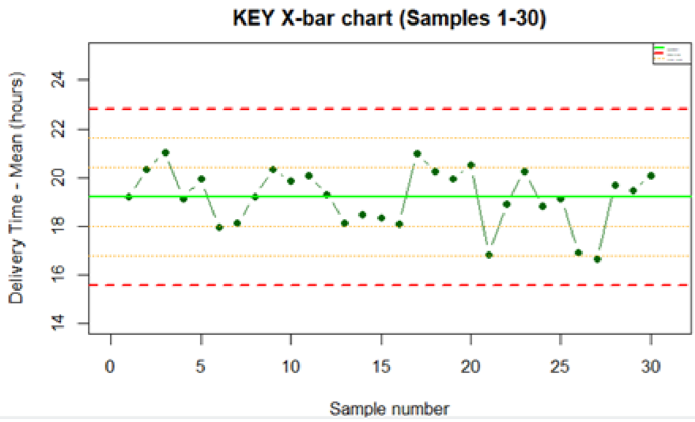

Figure 20: Keyboard X-bar chart for initial 30 samples

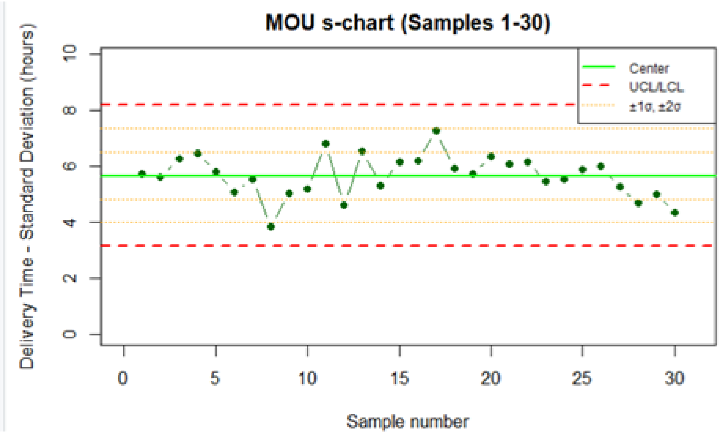

Figure 21: Mouse s-chart for initial 30 samples

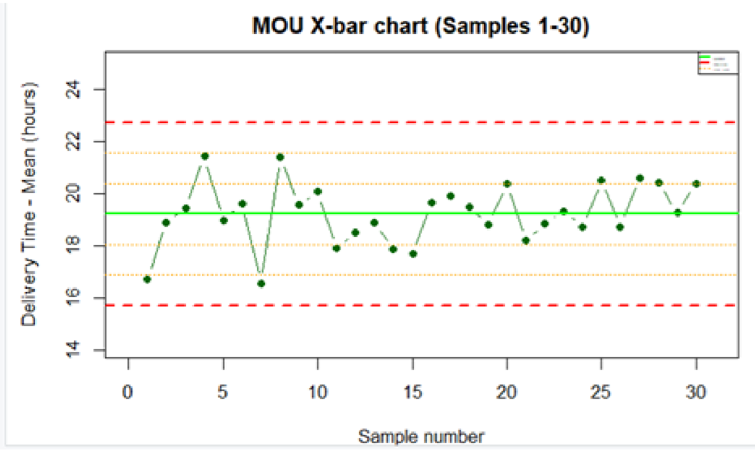

Figure 22: Mouse X-bar chart for initial 30 samples

The establishment of control limits using the initial 30 samples provided baseline parameters for subsequent process monitoring across all six product categories.

3.2.1 Monitor

The Monitor category demonstrated consistent variation patterns during the initial sampling period. The s-chart revealed standard deviation values clustering appropriately around the computed centreline, with no observations violating the established three-sigma boundaries. The corresponding X-bar chart showed process mean values fluctuating within expected ranges, though several points approached the upper two-sigma reference line. This behaviour suggested the process operated with inherent stability during baseline establishment, though slight upward drift in average delivery times warranted observation.

3.2.2 Laptop

Initial samples from the laptop category presented favourable process characteristics. Variation measurements on the s-chart remained well-distributed around the centreline without extreme departures, indicating stable dispersion during the baseline period. The X-bar chart similarly showed reasonable centring, with sample means generally maintaining position near the calculated centreline. Only occasional points approached the one-sigma boundaries, reflecting natural process variation rather than assignable cause influence. These patterns suggested the laptop delivery process possessed fundamental stability suitable for establishing reliable control parameters.

3.2.3 Cloud Subscription

Cloud Subscription exhibited exceptional control during initial sampling. The s-chart displayed remarkably tight clustering of variation measurements, with virtually all points remaining within one standard deviation of the centreline. This indicated unusually consistent process dispersion throughout the baseline period. The X-bar chart reinforced this assessment, showing process means maintaining proximity to target values with minimal wandering. The sustained stability suggested effective operational protocols and minimal external interference during the Phase I period, providing robust baseline parameters for ongoing monitoring.

3.2.4 Software

The Software category presented acceptable but less consistent baseline behaviour. Standard deviation measurements on the s-chart showed greater scatter compared to Cloud Subscription, though all points remained within established control limits. The X-bar chart revealed more substantial mean fluctuation, with several samples approaching two-sigma boundaries. While the process remained technically in statistical control during baseline establishment, the increased variability hinted at underlying instability that might manifest more prominently during extended monitoring.

3.2.5 Keyboard

Keyboard products demonstrated moderate baseline stability. The s-chart indicated acceptable variation control, though occasional measurements departed noticeably from the centreline. The X-bar chart showed process means drifting somewhat, with multiple samples positioned near or beyond one-sigma limits. These patterns suggested the presence of minor disturbances affecting delivery consistency during the initial phase, though no violations of three-sigma boundaries occurred to invalidate the baseline calculations.

3.2.6 Mouse

The Mouse category's baseline data revealed early warning signs of process instability. The s-chart showed wider scatter in variation measurements compared to more stable categories, indicating inconsistent process dispersion even during the controlled baseline period. The X-bar chart displayed substantial mean fluctuation, with numerous samples approaching or touching two-sigma boundaries. While control limits could be established from these initial samples, the patterns foreshadowed the more severe instability that would emerge during extended monitoring.

3.3 Control Chart Monitoring for all samples

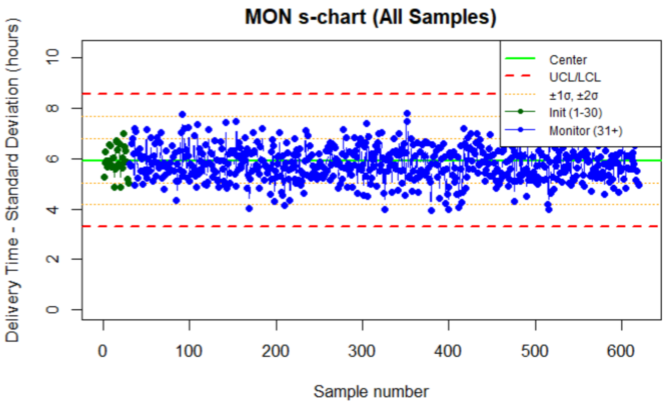

Figure 23: Monitor s-chart for all samples

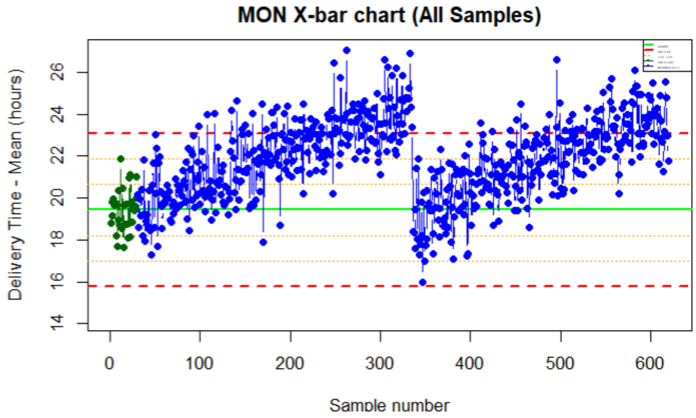

Figure 24: Monitor X-bar chart for all samples

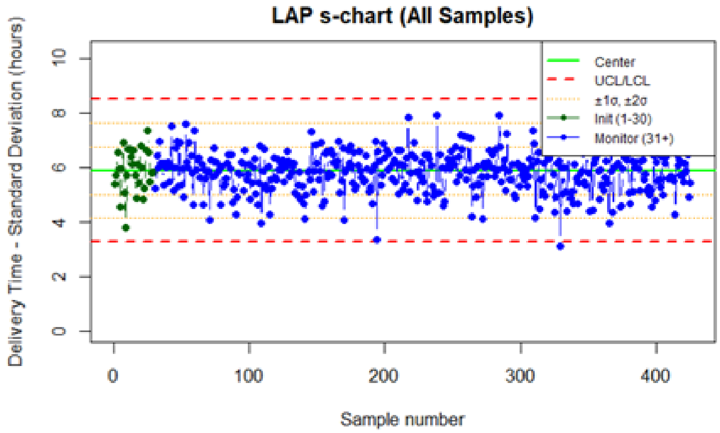

Figure 25: Laptop s-chart for all samples

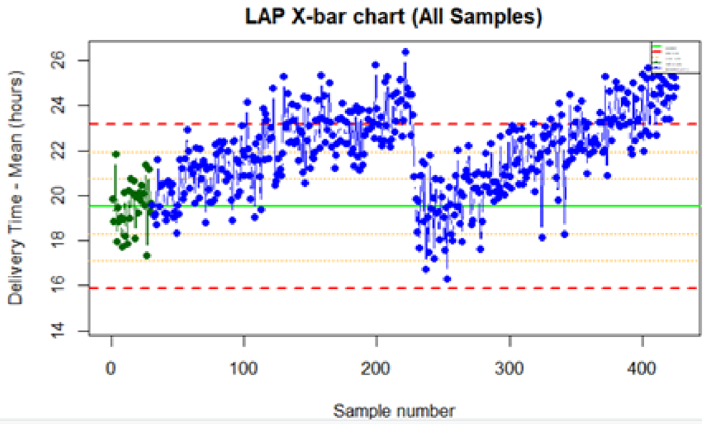

Figure 26: Laptop X-bar chart for all samples

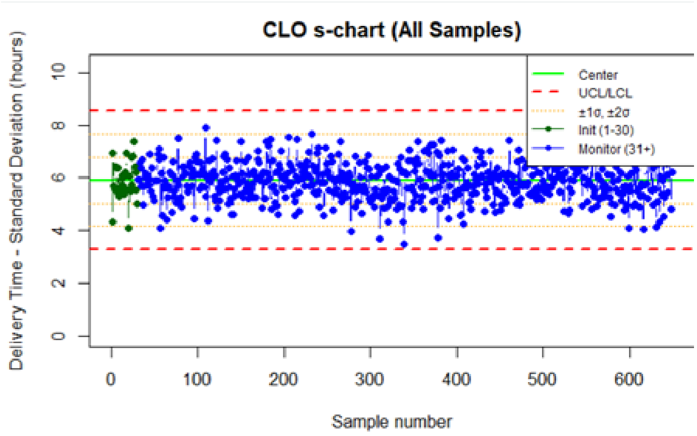

Figure 27: Cloud Subscription s-chart for all samples

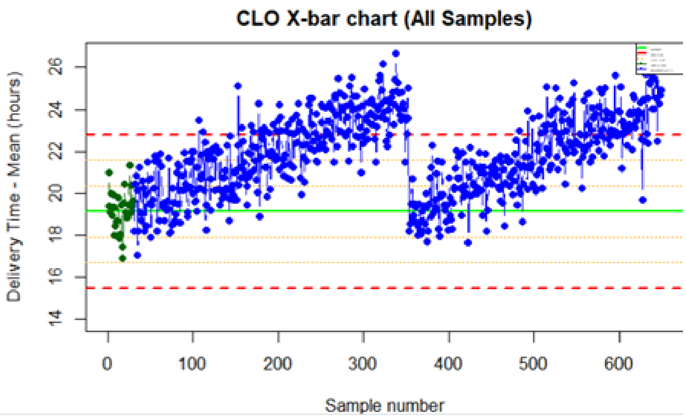

Figure 28: Cloud Subscription X-bar chart for all samples

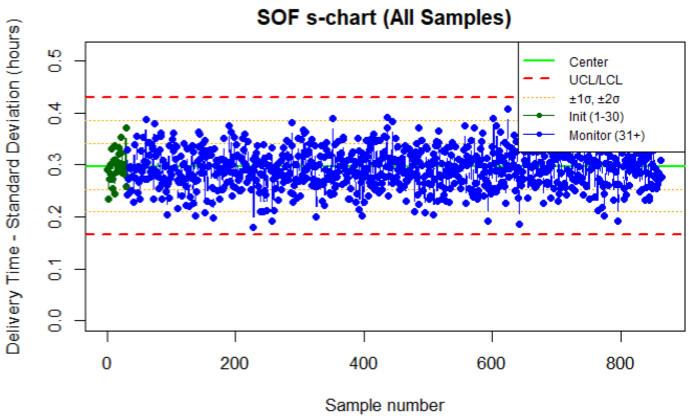

Figure 29: Software s-chart for all samples

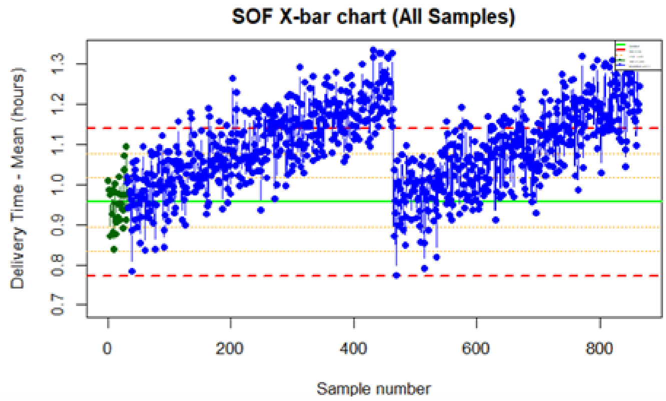

Figure 30: Software X-chart for all samples

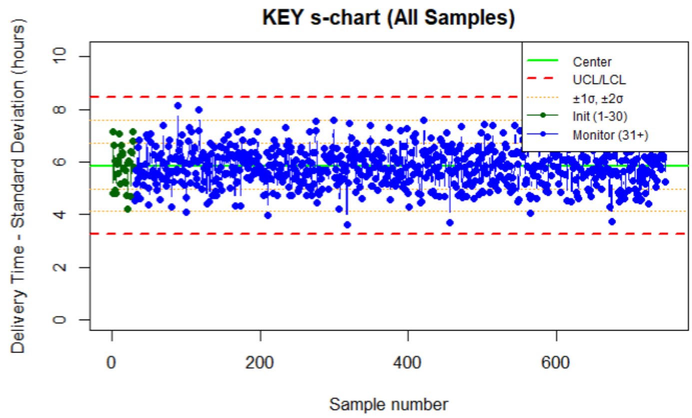

Figure 31: Keyboard s-chart for all samples

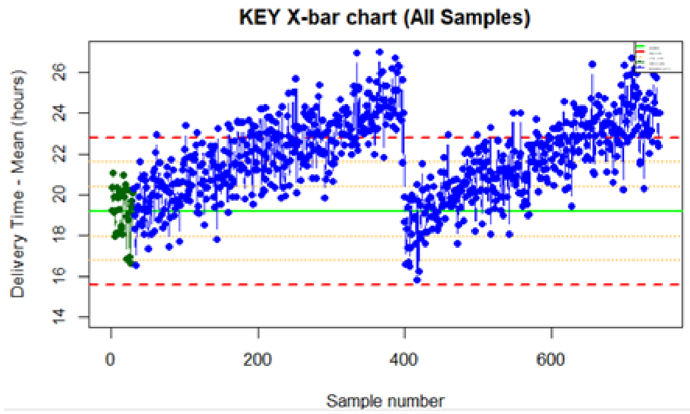

Figure 32: Keyboard X-bar chart for all samples

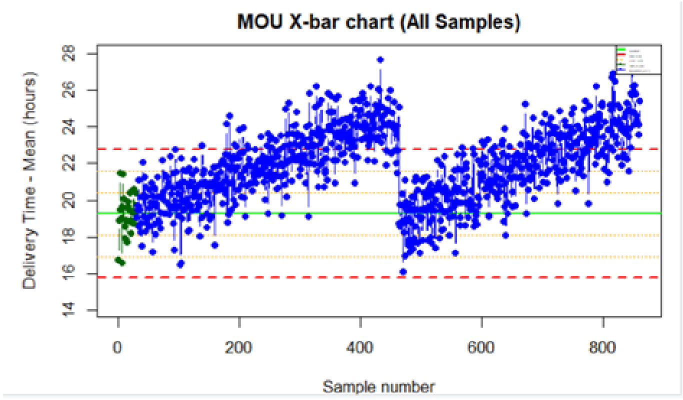

Figure 33: Mouse X-bar chart for all samples

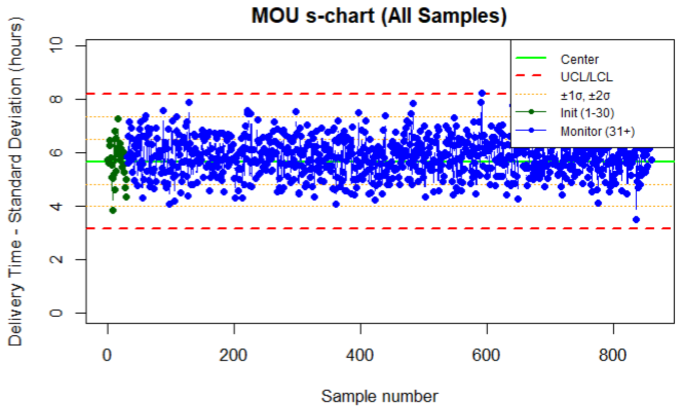

Figure 34: Mouse s-chart for all samples

3.3.1 Monitor

Monitor category performance during extended monitoring confirmed the moderately stable characteristics observed during baseline establishment. The s-chart continued showing acceptable variation control throughout the monitoring period, with standard deviation measurements remaining generally within expected boundaries. However, the X-bar chart revealed increasing frequency of points exceeding two-sigma limits as monitoring progressed. Analysis identified 157 instances where four consecutive sample means exceeded the upper two-sigma boundary, violating the detection rules. These violations clustered in specific regions, with early occurrences around samples 134, 179, and 190 and later manifestations near samples 605, 610 and 615. Despite these concerning patterns, the process maintained 49 consecutive samples between plus and minus one-sigma limits during one stable period, demonstrating capability for sustained control when conditions remained favourable.

3.3.2 Laptop

Laptop products maintained relatively robust process control throughout extended monitoring. The s-chart showed continued stability in variation measurements, with few dramatic departures from expected patterns. The X-bar chart demonstrated the fewest control rule violations among all categories examined, recording only 123 instances of four consecutive points beyond two-sigma limits. The first three violations occurred at samples 119, 130 and 131, while the final three appeared at samples 420, 421 and 422. Additionally, the process achieved 36 consecutive samples remaining between one-sigma boundaries, indicating periods of exceptional stability. This performance suggested inherent process robustness or particularly effective operational management compared to other product categories.

3.3.3 Cloud Subscription

Cloud Subscription delivery processes validated the exceptional stability observed during initial sampling. Throughout extended monitoring, the s-chart maintained tight control over variation, with standard deviation measurements rarely departing substantially from the centreline. The X-bar chart demonstrated sustained process centring, achieving 36 consecutive samples between one-sigma limits, the longest stable run among all categories analysed. Despite this impressive stability, the category recorded 203 violations of the four-consecutive-points rule, with first occurrences at samples 122, 179 and 180, and final violations at samples 644, 645 and 646. The combination of extended stable periods with clustered violations suggested the process operated under tight control most of the time, with occasional disturbances creating temporary instability that subsequently resolved.

3.3.4 Software

Software category extended monitoring confirmed the instability tendencies suggested during baseline establishment. The s-chart showed increasing scatter in variation measurements as monitoring continued, indicating progressively less consistent process dispersion. The X-bar chart revealed severe control problems, recording 259 violations of the consecutive-points rule, the second-highest frequency among all categories. Initial violations appeared at samples 202, 237 and 244, with final violations occurring at samples 859, 860 and 861, demonstrating problems persisting throughout the entire monitoring period. The process managed only 39 consecutive samples between one-sigma limits during its longest stable run, substantially shorter than better-controlled categories. These patterns indicated persistent assignable causes affecting delivery performance that remained unidentified and unresolved throughout the observation period.

3.3.5 Keyboard

Keyboard products exhibited moderate instability during extended monitoring. The s-chart revealed periodic spikes in variation measurements, suggesting intermittent process disturbances affecting delivery consistency. The X-bar chart recorded 219 violations of control rules, with initial occurrences at samples 112, 113 and 114, and final violations at samples 741, 742 and 743. The process achieved only 32 consecutive samples between one-sigma limits during its longest stable period, indicating frequent interruption of stable operation by assignable cause influences. The patterns suggest that external factors occasionally disrupted deliveries, leading to brief periods of stability followed by repeated disturbances.

3.3.6 Mouse

Mouse category extended monitoring revealed the most severe process control deficiencies among all products examined. The s-chart identified sample 592 exceeding the upper three-sigma control limit, representing a statistically significant increase in process variation requiring immediate investigation. This violation indicated fundamental process instability beyond normal operational fluctuation. The X-bar chart demonstrated the highest violation frequency, recording 255 instances of four consecutive points beyond two-sigma limits. Initial violations occurred at samples 194, 235 and 236, with final violations at samples 855, 856 and 857, indicating chronic instability affecting nearly the entire monitoring period. The process achieved only 21 consecutive samples between one-sigma limits during its longest stable run, demonstrating inability to maintain consistent control for extended periods. These severe patterns indicated multiple unresolved assignable causes fundamentally compromising delivery process capability and requiring urgent corrective intervention to restore acceptable performance levels.

3.4 Process capability analysis

Control charts were developed using delivery time data, with 24 observations per sample. The first 30 samples were used to establish control limits (centerline and ±1σ, ±2σ, ±3σ boundaries) for effective process monitoring.

Process capability indices were calculated using the following formulas:

$Cp = (USL - LSL) / 6\sigma$

$Cpu = (USL - \mu) / 3\sigma$

$Cpl = (\mu - LSL) / 3\sigma$

$Cpk = \min(Cpu, Cpl)$

Assuming a lower specification limit (LSL) of 0 hours and an upper specification limit (USL) of 32 hours, the results show that all product categories have Cp and Cpk values below 1.33, indicating that the processes cannot consistently meet customer delivery expectations. As shown in the table, the Mouse (Cpk = 0.727), Keyboard (Cpk = 0.729), Cloud Subscription (Cpk = 0.717), Laptop (Cpk = 0.696), and Monitor (Cpk = 0.700) categories are not capable. The Software category shows Cp = 18.135 and Cpk = 1.083, suggesting higher capability, though the Cp value appears unusually large and may indicate a measurement or scaling irregularity.

A process is considered capable when Cpk ≥ 1.33; therefore, these results confirm excessive variability and poor centring across all 6 product types. The Mouse and Software categories also showed frequent control chart violations, while Cloud Subscription and Monitor were relatively stable. To improve process capability, variability reduction should be prioritized through refined standard operating procedures (SOPs), consistent operator training, and improved process automation.

ProductType	N_obs	Mean	StdDev	Cp	Cpu	Cpl	Cpk	Capable
MOU	1000	19.298	5.828	0.915	0.727	1.104	0.727	NO
KEY	1000	19.276	5.815	0.917	0.729	1.105	0.729	NO
SOF	1000	0.955	0.294	18.135	35.188	1.083	1.083	NO
CLO	1000	19.226	5.941	0.898	0.717	1.079	0.717	NO
LAP	1000	19.606	5.934	0.899	0.696	1.101	0.696	NO
MON	1000	19.410	5.999	0.889	0.700	1.079	0.700	NO

Table 1: Results from R for process capability analysis

3.4.1 Rule A

Only the Mouse category recorded a sample exceeding the upper three-sigma control limit at sample 562, indicating a statistically significant increase in process variation.

3.4.2 Rule B

Most consecutive s-chart samples remaining between ±1σ limits:

- Cloud Subscription: 36 consecutive samples
- Monitor: 49 consecutive samples
- Laptop: 36 consecutive samples
- Software: 39 consecutive samples
- Mouse: 21 consecutive samples
- Keyboard: 32 consecutive samples

Processes such as Monitor and Cloud Subscription exhibit strong stability, maintaining extended sequences of samples within the ±1σ range, reflecting effective process management and minimal variation. In contrast, shorter stable runs for Mouse and Keyboard suggest greater process variability and potential exposure to external or operational disturbances, warranting closer monitoring and refinement.

3.4.3 Rule C

Four consecutive X-bar samples exceeding upper 2σ limits:

- **Monitor:** First three violations at samples 134, 179, 190; last three at 605, 610, 615; total occurrences: 157
- **Software:** First three violations at samples 202, 237, 244; last three at 859, 860, 861; total occurrences: 259
- **Keyboard:** First three violations at samples 112, 113, 114; last three at 741, 742, 743; total occurrences: 219
- **Cloud Subscription:** First three violations at samples 122, 179, 180; last three at 644, 645, 646; total occurrences: 203
- **Mouse:** First three violations at samples 194, 235, 236; last three at 855, 856, 857; total occurrences: 255
- **Laptop:** First three violations at samples 119, 130, 131; last three at 420, 421, 422; total occurrences: 123

The highest frequency of out-of-control signals was observed for Software and Mouse, indicating process instability and the need for targeted corrective measures. Laptop, on the other hand, recorded the fewest violations, suggesting a well-controlled and reliable process. These findings imply that best practices from the laptop production process could be leveraged to strengthen control in higher-variation categories such as Mouse and Software.

Part 4: Type I and Type II Error Estimation

4.1 Methodology

Erroneous category labels were corrected using conditional mapping functions (mutate() and case_when()). The probability of false alarms (Type I error, α) and missed detections (Type II error, β) were derived using R's pnorm() and qnorm() functions under the standard normal distribution. Scenarios simulating out-of-control shifts were modelled to estimate detection sensitivity and chart robustness. This approach connected theoretical quality-control principles to the actual reliability of process-monitoring mechanisms, ensuring that subsequent control and capability assessments were statistically valid.

4.2 Type I error

This represents the probability of concluding that a process is out of control when it is in control.

- **Rule A:** $\alpha \approx 0.0027$ per observation.
- **Rule B:** Detects unusual process stability by identifying k consecutive samples all falling within ±1σ limits. The probability that a single sample falls within ±1σ equals 0.6827 under normal distribution. For k consecutive samples, this probability equals (0.6827)^k, making the Type I error α = 1 − (0.6827)^k. For k = 10 consecutive samples, α = 1 − (0.6827)^10 ≈ 1 − 0.0188 ≈ 0.981, indicating 98.1% probability of observing this pattern when the process is in control.
- **Rule C:** $\alpha \approx 2.68\times10^{-7}$ (extremely low).

Thus, Rule C is the most conservative, minimizing false alarms.

4.3 Type II error

The following calculations estimate the likelihood of making type II (Consumer's) Errors for a bottle filling process, that should be centred on 25.05 litres as the process average and CL of an xbar chart and has an UCL of 25.089 and LCL of 25.011 litres. Assuming the process has shifted, the process has moved to an average fill volume of 25.028 litres and now has a xbar standard deviation of 0.017 instead of 0.013 litres. In the type II error, the Ha is true, but we fail to identify this, due to the sample xbar and s values being between LCL and UCL of each chart.

Type II Error (β):
Given: μ_0 = 25.05 L, μ_1 = 25.028 L, σ = 0.017, LCL = 25.011, UCL = 25.089

$$Z_L = \frac{25.011 - 25.028}{0.017} = -1.0, Z_U = \frac{25.089 - 25.028}{0.017} = 3.59$$
$$\beta = \Phi(Z_U) - \Phi(Z_L) = \Phi(3.59) - \Phi(-1.0) \approx 0.9998 - 0.1587 = 0.8411$$

Hence, β ≈ 0.84, meaning there is an 84% chance of missing a true process shift, indicating low detection sensitivity

4.3 Data Correction and Re-Analysis

4.3.1 Data Correction Methodology

The following updates were implemented to address errors in the products_Headoffice2025.csv file and enhance the products_data.csv file. Initially, corrections will be applied to the product types listed from lines 11 to 60 for each category within the products_Headoffice2025.csv file. For instance, entries such as "NA011 Software Alice blue silk 823.51 14.59" and "SOF010 Software cornflower blue matt 399.43 17.08" contain inaccuracies, where the green values are correct as stated in the final report file, but the yellow values (representing prices and markups) are erroneous. These will be revised to align with the accurate data, such as "SOF010 Software cornflower blue matt 396.72 23.47" and "SOF011 Software Alice blue silk 511.53 25.05," utilizing the correct selling prices and markups sourced from the products_data.csv file. This correction will be consistently applied across all affected product types throughout the document.

Subsequently, any product IDs beginning with "NA" (e.g., "NA011") will be amended to reflect the appropriate prefix, such as "SOF" for Software or "Key" for Keyboard, based on the respective product category. Additionally, the selling prices and markups in the products_Headoffice2025.csv file, which repeat every 10 items (e.g., items 1, 11, 21, etc., and items 2, 12, 22, etc.) due to their association with the same brand and model, will be updated. These values will be replaced with the correct set of 10 values derived from the products_data.csv file, ensuring uniformity and accuracy across the sequence.

Furthermore, the products_data.csv file will be modified by aligning the Category column with the corresponding ProductID column to ensure accurate categorization for each product. Upon completion of these revisions, the updated products_data.csv file will be saved as a new file named products_data2025.csv for future use and reference.

4.3.2 Impact on Descriptive Statistics

The graph remains unchanged before and after the update. Although the Product table's Category column was modified to correctly match the ProductID, this adjustment did not affect the overall count, as shown in the graph below.

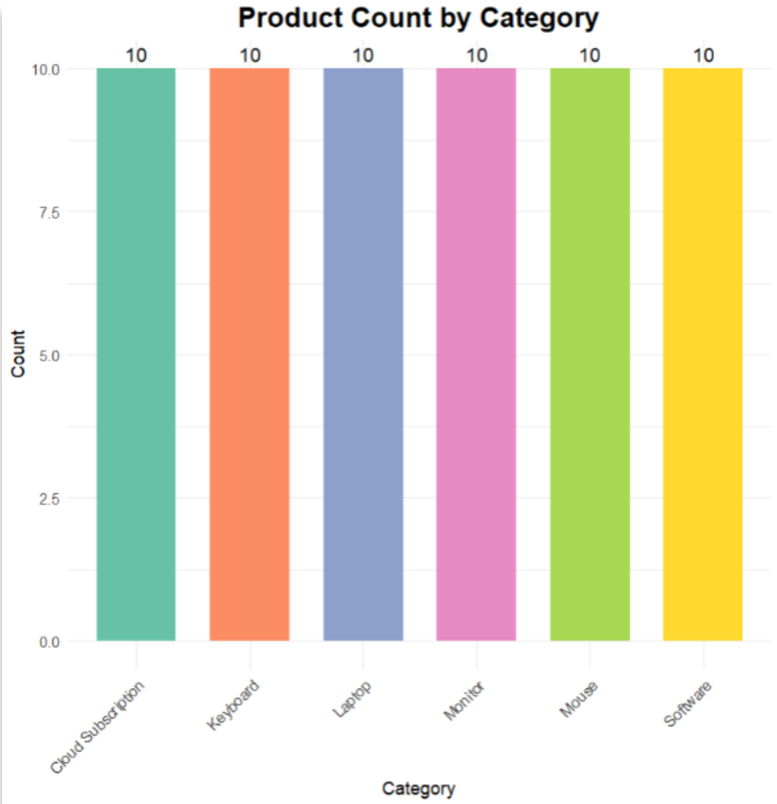

Figure 35: Product count by category updated

The updated graph displays different average selling prices per category because the Category field was aligned with the Product ID. In the updated Product2025 version, the product SOF002 was correctly reclassified as Software instead of Cloud Subscription, with a selling price of 505.26 and a markup value of 10.43. This correction significantly impacted the results.

The laptop category's average selling price increased from 5,217.54 to 10,886.43, while the Monitor category showed a modest rise from 5,014.17 to 6,310.52. In contrast, Cloud Subscription, Keyboard, Software and Mouse categories experienced notable decreases. For instance, the Keyboard dropped from 4,638.17 to 644.66, and the Mouse from 4,585.46 to 394.70.

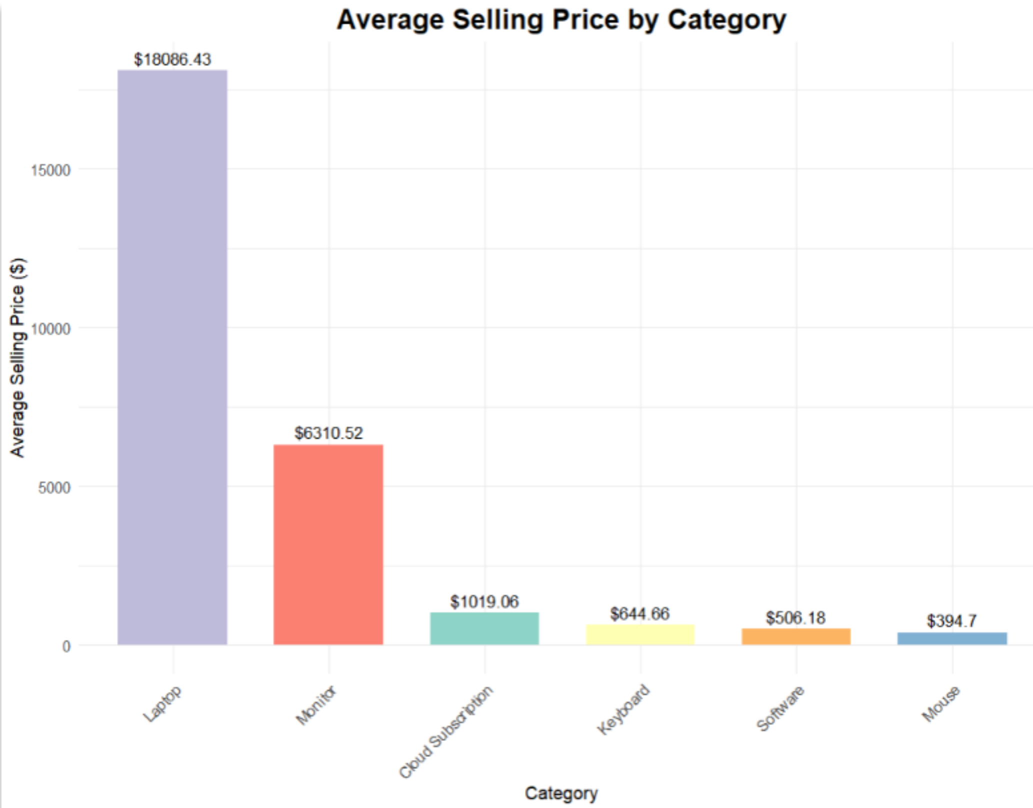

Figure 36: Average selling price by category updated

The main difference is observed in the Keyboard category, where the average markup value increased significantly from 20.16 to 23.98, making it the highest among all categories. Both Keyboard and Monitor categories show an increase in average markup values, while Mouse, Laptop, Cloud Subscription and Software categories all experienced a decrease.

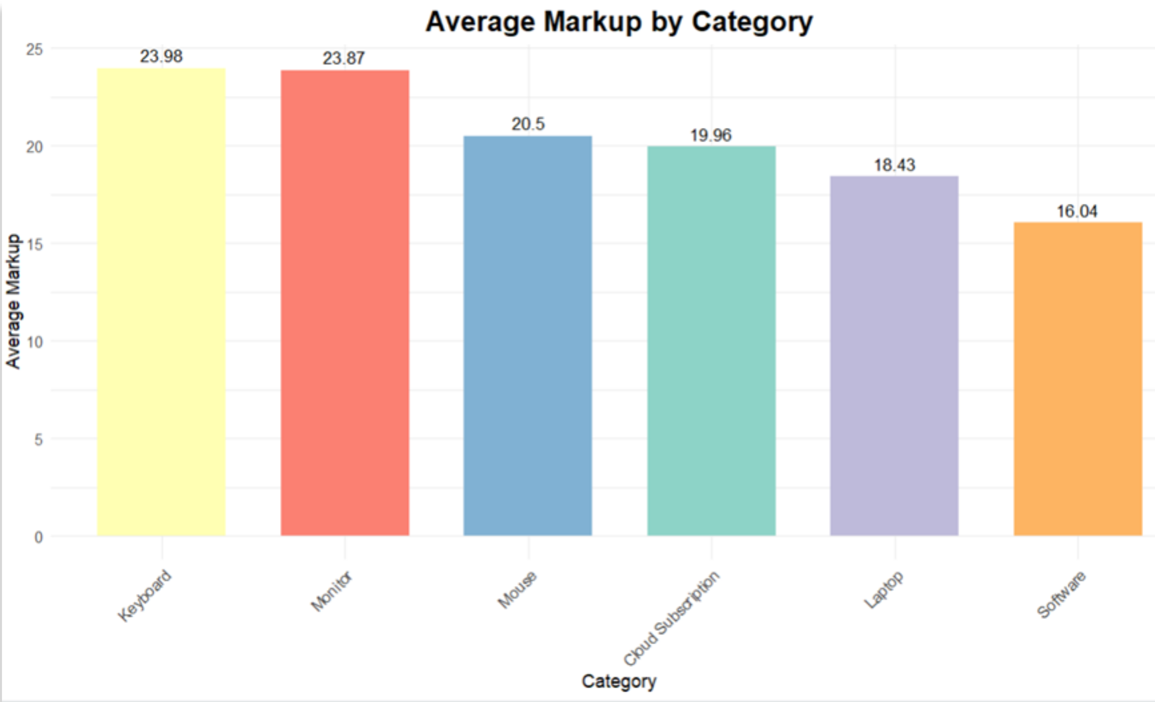

Figure 37: Average Markup by Category updated

The rest of the graphs in the first questions remain unchanged.

4.3.3 2023 Sales Value

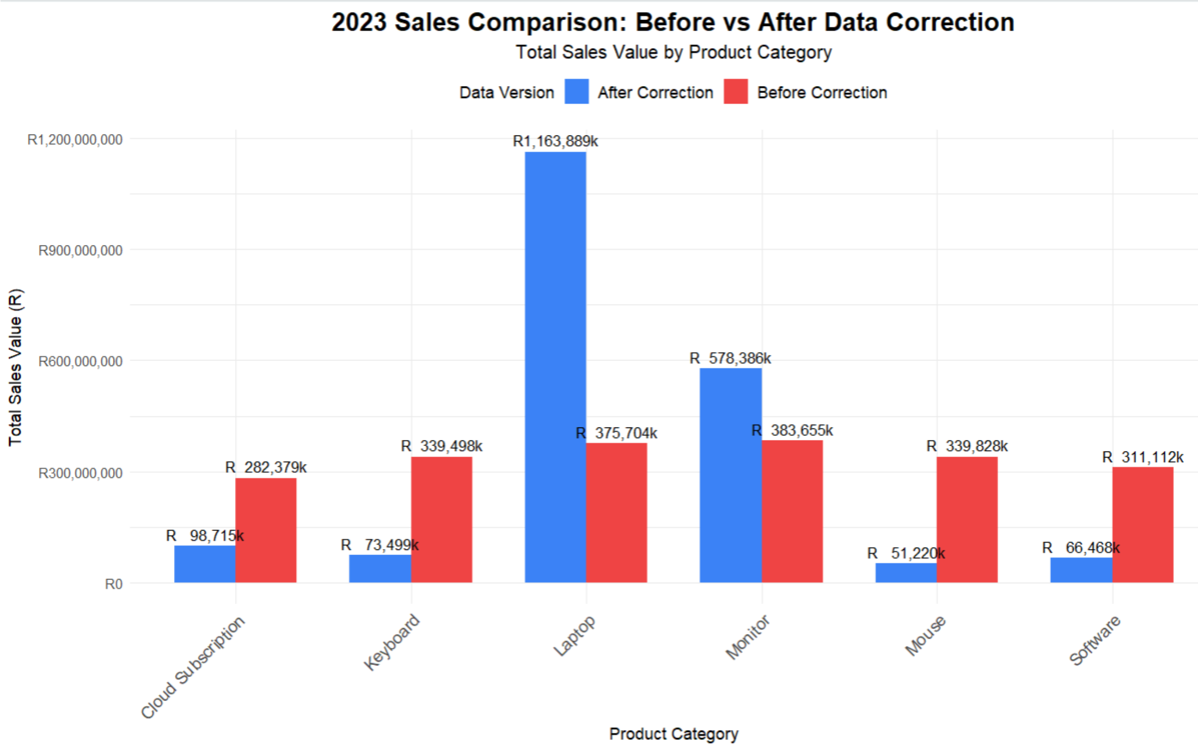

Figure 38: Total Sales Value by Product Category

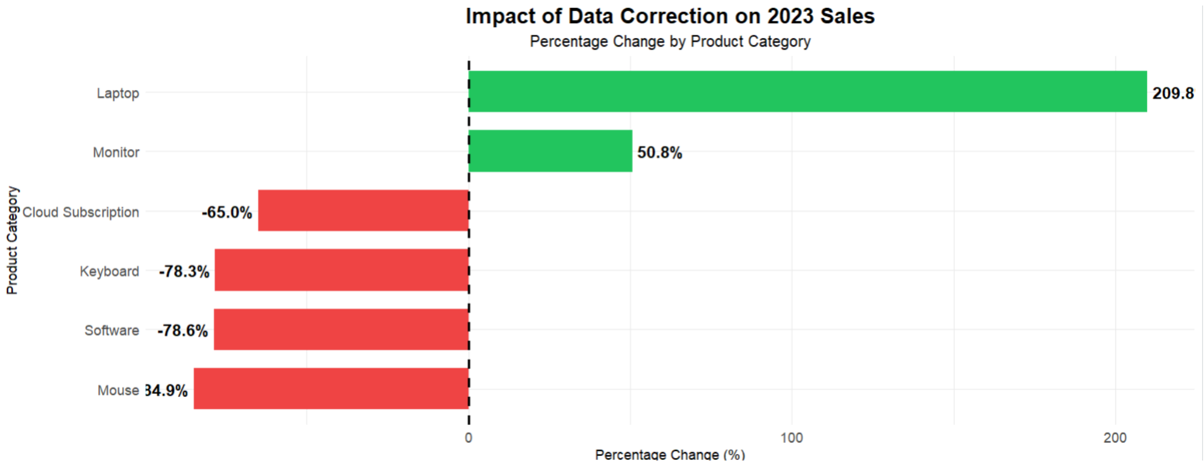

Figure 39: Impact of Data Correction on 2023 sales

Figure 38 illustrates the comparison of 2023 sales values before and after the data correction. The most significant adjustment occurred in the laptop category, where sales rose sharply from R375.7 million to R1,163.9 million, representing a 209.8% increase. This correction revealed that laptop prices had been considerably undervalued in the original dataset. Similarly, monitor sales increased from R383.7 million to R578.4 million, reflecting a 50.8% rise and indicating previous under-pricing. In contrast, several categories experienced substantial decreases following the correction. Mouse sales fell by 84.9% (from R339.8 million to R51.2 million), Keyboard by 78.3% (R339.5 million to R73.5 million), Software by 78.6% (R311.1 million to R66.5 million), and Cloud Subscription by 65.0% (R282.4 million to R98.7 million).

Figure 39 presents these percentage changes visually, clearly distinguishing between the categories that benefited from correction and those that were overvalued. The increases observed in the Laptop and Monitor categories confirm that their prices were previously understated, while the sharp declines in Mouse, Keyboard, and Software reveal significant overpricing in the original data. After correction, Laptops now account for approximately 57.3% of total 2023 revenue, establishing them as the company's leading income source. Meanwhile, the reduced figures for peripheral products demonstrate that their contribution to overall revenue had been substantially overestimated prior to the data adjustment.

Part 5: Profit Optimisation

5.1 Methodology

This analysis examines the relationship between staffing levels and business performance at a coffee shop using one year of recorded service time data. The objective was to determine the optimal number of baristas required to maximize monthly profitability while maintaining service quality standards. The business operates eight hours daily and generates material profit of R30 per customer, excluding labour costs. Each barista costs R1,000 per day (R30,000 monthly), with operational requirements mandating a minimum of two baristas on duty.

5.2 Analysis of Shop 1

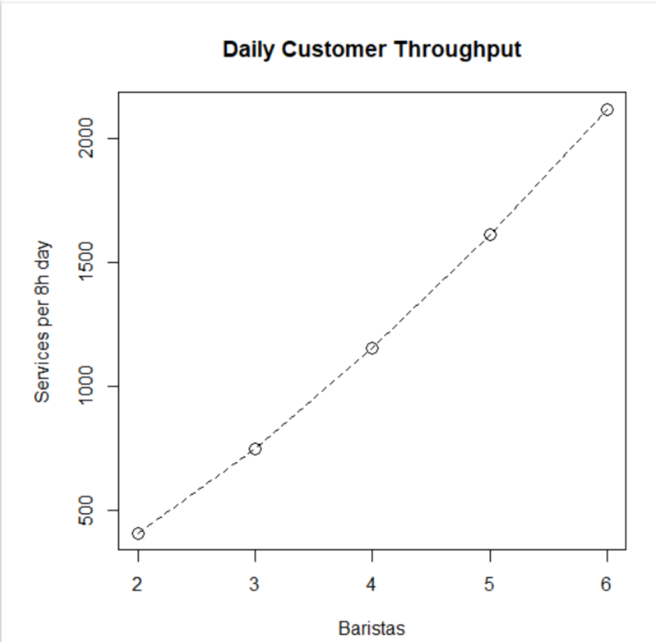

Figure 40: Daily customer throughput for timeToServe

Figure 40 shows the relationship between staffing and daily customer capacity. Throughput was calculated as Daily Capacity = Operating Minutes / Average Service Time, where operating minutes equal 480 (8 hours × 60 minutes). With two baristas, daily capacity equals 407 customers. This expands to 748 customers (3 baristas), 1,152 customers (4 baristas), 1,610 customers (5 baristas), and 2,117 customers (6 baristas). The 420% capacity increase from minimum to optimal staffing demonstrates strong economies of scale in service delivery, where faster service times combine multiplicatively with additional service positions to generate exponential throughput growth.

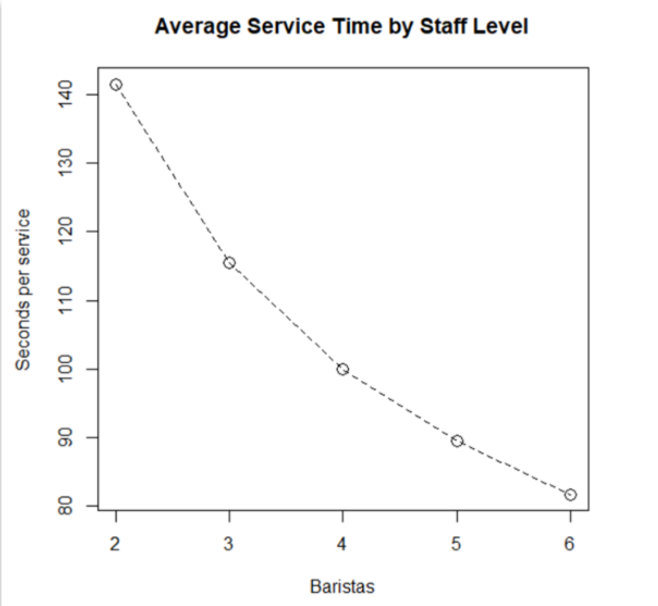

Figure 41 :Average service time by staff level for timeToServe

Figure 41 demonstrates an inverse relationship between staffing levels and average service time per customer. The relationship follows a logarithmic decay pattern characteristic of capacity expansion with diminishing marginal returns. At minimum staffing (2 baristas), average service time reaches 142 seconds per customer. Increasing to three baristas reduces service time to 115 seconds, representing a 19% improvement. Further staffing increases yield progressively smaller gains: four baristas achieve 100 seconds (13% improvement), five baristas reach 89 seconds (11% improvement), and six baristas deliver 82 seconds (8% improvement). This diminishing return pattern reflects the operational reality that each additional staff member contributes less incremental benefit due to spatial constraints, coordination overhead, and task interdependencies within the service environment.

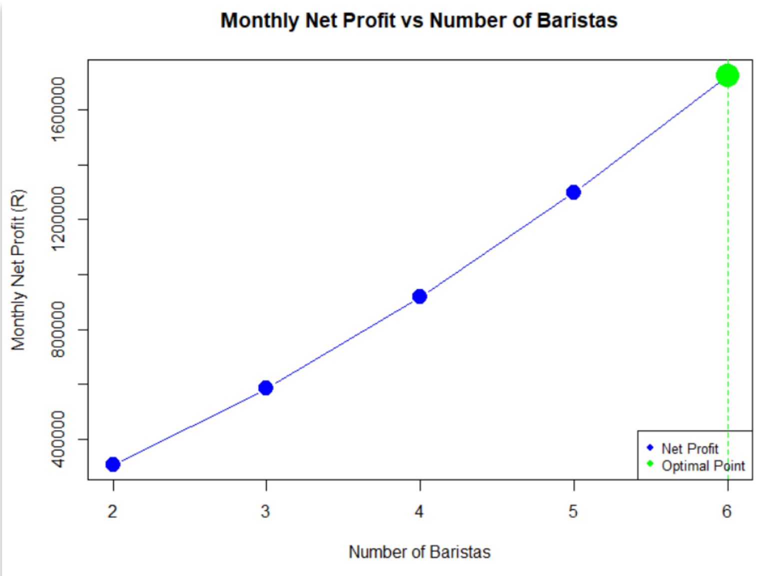

Figure 42: Monthly Net profit per number of baristas for timeToServe

Figure 42 presents the profit optimization analysis across staffing configurations. Monthly net profit was calculated as (Daily Customers × 30 days × R30 profit margin) − (Number of Baristas × R30,000 monthly salary). Results indicate monotonically increasing profitability across the examined range: two baristas generate R306,300 monthly profit, three baristas yield R584,400, four baristas produce R916,800, five baristas achieve R1,299,000, and six baristas maximize profit at R1,725,300 monthly.

Marginal analysis reveals increasing returns to scale throughout the staffing range. The third barista adds R278,100 profit against R30,000 cost (9.3× return), the fourth adds R332,400 (11.1× return), the fifth contributes R382,200 (12.7× return), and the sixth delivers R426,300 (14.2× return). This increasing marginal return pattern indicates that the operation has not yet reached capacity saturation, suggesting profitability will likely continue increasing beyond six baristas if demand permits.

The optimal configuration of six baristas delivers monthly revenue of R1,905,300 from serving 63,496 customers over 30 operating days, with labour costs of R180,000, yielding net profit of R1,725,300. Average service time of 82 seconds ensures customer satisfaction while maximizing throughput. Compared to minimum staffing, the six-barista configuration costs an additional R120,000 monthly in labour but generates R1,418,563 additional profit, representing an 1183% return on incremental investment.

5.3 Analysis of Shop 2

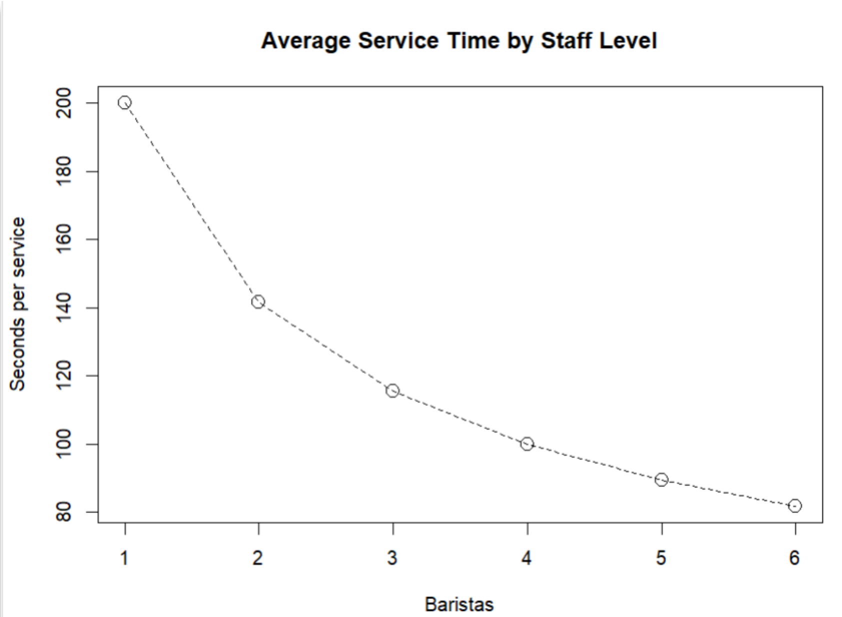

Figure 43: Average service time by staff level for timeToServe2

Figure 43 demonstrates a powerful inverse relationship between staffing levels and the time required to serve each customer. The data reveals a dramatic improvement in service speed as more baristas are added to the team. Starting with just 1 barista working alone, the average service time is approximately 200 seconds per customer, which represents extremely slow service that would likely lead to very long queues, customer complaints, and lost sales as frustrated customers leave before being served. When staffing increases to 2 baristas, service time drops significantly to about 142 seconds, showing the immediate benefit of having even one additional team member. The improvements continue progressively: 3 baristas achieve approximately 115 seconds, 4 baristas reach 100 seconds, 5 baristas deliver 89 seconds, and 6 baristas provide the fastest service at just 82 seconds per customer. The curve shows diminishing returns, meaning that each additional barista produces smaller incremental improvements in service speed, but the overall trend is clear and consistent. This pattern demonstrates how additional staff enables better workflow distribution, reduces bottlenecks, allows for task specialization, and provides the capacity to handle multiple customers simultaneously, all of which contribute to faster, more efficient service delivery

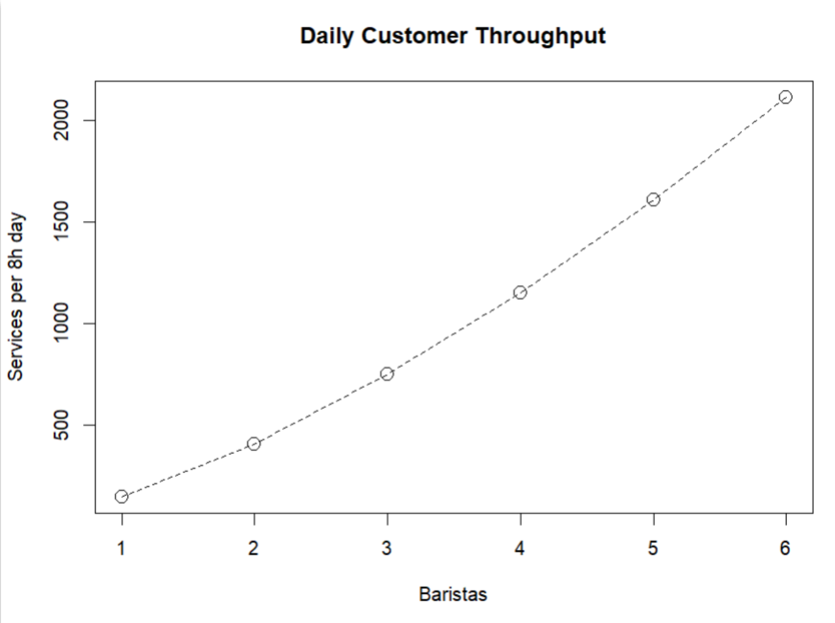

Figure 44: Daily customer throughput for timeToServe2

Figure 44 illustrates the transformative impact that staffing levels have on the coffee shop's daily capacity to serve customers during its 8-hour operating day. This graph shows a nearly linear upward trend, indicating that customer throughput grows consistently and predictably with each additional barista hired. With only 1 barista working, the shop can serve approximately 144 customers per day, which represents severely limited capacity. Adding a second barista nearly triples this capacity to about 407 customers per day, demonstrating the exponential gains from moving away from single-person operations. The growth continues impressively: 3 baristas enable serving 748 customers, 4 baristas handle 1,152 customers, 5 baristas manage 1,610 customers and 6 baristas can serve an 2,117 customers per day. This increase in throughput occurs because faster service times and more available staff members. Each barista can process more customers when service is faster, and having multiple baristas working simultaneously multiplies this capacity even further.

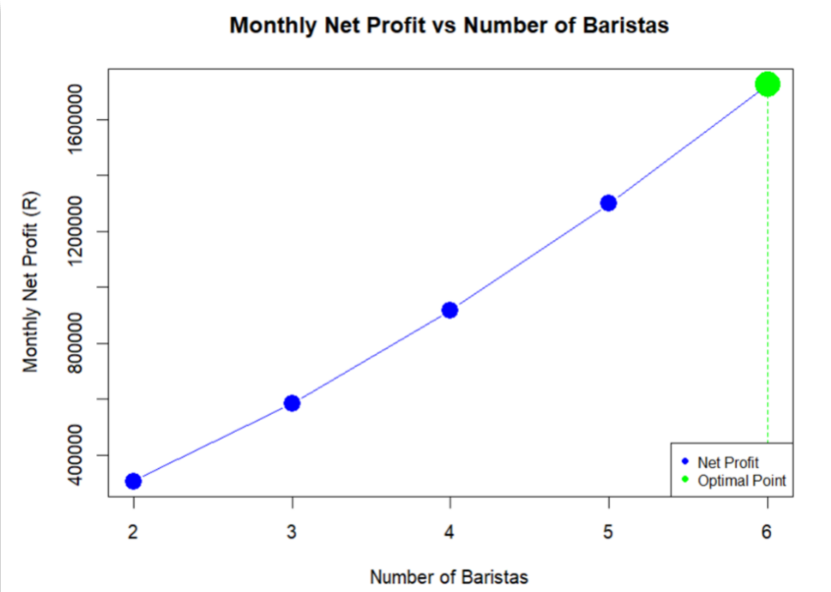

Figure 45: Monthly net profit vs number of baristas for TimeServe2

Figure 45 reveals the financial implications of different staffing strategies by plotting monthly net profit against the number of baristas employed. This graph tells the complete financial story and justifies the optimal staffing decision.

Starting with 1 barista, the business generates approximately R99,500 in monthly profit. This is profitable but is severely limited by capacity constraints. Adding 1 barista increases profit substantially to R306,000, more than tripling profitability despite doubling labour costs. The upward trend continues consistently: 3 baristas yield R584,000, 4 baristas produce R917,000, 5 baristas generate R1,299,000, and 6 baristas maximize profit at an impressive R1,725,000 per month. The green dot marking the optimal point at 6 baristas, combined with the vertical green reference line, clearly identifies this as the profit-maximizing staffing level.

Profit continues to increase as more baristas are hired, even though labour costs also increase. This is because the extra income from serving more customers is more than the added cost of another employee. Each barista costs R30,000 per month but allows hundreds of extra sales, each earning R30 in profit. This shows that, under current prices and demand, adding staff directly improves profitability.

The ideal staffing level is six baristas, which gives the highest profit and best performance overall. At this level, monthly revenue is about R1,905,000 from 63,496 customers over 30 days. Labour costs total R180,000, resulting in a net profit of around R1,725,000. The average service time is 82 seconds per customer, and the shop can serve about 2,117 customers a day, ensuring fast service and high customer satisfaction.

Compared to two baristas, increasing to six adds R120,000 in labour costs but generates an extra R1,418,563 in monthly profit. This makes six baristas the ideal balance between cost, efficiency and profit. The business should maintain this staffing level and track demand in case more staff could still increase profit, though further gains may eventually slow as efficiency improvements level off.

<u>5.4 Comparison and Interpretation:</u>

The analysis confirms that adding baristas substantially improves throughput, reduces service time, and increases profitability in both shops; however, the magnitude and rate of improvement differ between the two datasets.

In Shop 2 (timeToServe2), the single-barista operation performs poorly, with a mean service time of 200.16894 seconds and only 143.8785 customers per day. When a second barista is added, performance improves dramatically: service time drops to 141.51462 seconds, and daily throughput rises to 407.0251 customers—a 183% increase in customers served and a 29% reduction in service time. This early surge highlights how severely under-staffed the single-barista setup was, with the second employee immediately removing a major process bottleneck. Further additions continue to improve efficiency, but at a decreasing rate, as service time falls to 81.64272 seconds and throughput increases to 2 116.5390 customers per day with six baristas.

For Shop 1 (timeToServe), which begins with two baristas as its minimum staffing level, the pattern is smoother and less extreme. Average service time decreases from 100.17098 seconds at two baristas to 33.35565 seconds at six, while customer throughput rises from 575.0168 to 5 180.5322 customers per day—an almost nine-fold increase. Although substantial, the improvement is more gradual than in Shop 2 because the operation was already functioning above its critical staffing threshold.

Profitability mirrors these operational dynamics. In Shop 2, monthly net profit rises from R 99 490.62 with one barista to R 1 724 885.06 with six, a 17.34-fold increase. The relative gradient between the first and sixth baristas,

$$(1724885.06–99490.62)/99490.62 = 16.34,$$

confirms exceptionally high returns in the lower staffing range. Shop 1, meanwhile, shows a larger absolute profit but a lower proportional gain, increasing from R 457 515.16 at two baristas to R 4 482 478.98 at six. The profit ratio of 9.80 demonstrates steady but less volatile growth, as efficiency gains accumulate rather than surge.

Comparatively, Shop 2 exhibits higher labour sensitivity: early staffing increments yield large efficiency and profit gains, showing that the operation was initially constrained by inadequate capacity. Shop 1 displays operational maturity, where additional staff improve performance but with diminishing marginal returns due to coordination limits and space utilisation.

Both models reach their maximum profit at six baristas, yet for different reasons. Shop 2's optimum represents recovery from under-staffing—profit growth is demand-driven and capacity-constrained—whereas Shop 1's optimum reflects process refinement within a stable, already efficient system. In practical terms, Shop 2 should prioritise maintaining sufficient baseline staffing to avoid service bottlenecks, while Shop 1 should focus on workflow optimisation, task sequencing, and queue management to sustain profitability without unnecessary labour expansion.

Part 6: Analysis of Variance (ANOVA) for Delivery Times:

6.1 Methodology

A one-way Analysis of Variance (ANOVA) was performed in R to test whether mean delivery times differed across twelve months. The calculated F-statistic, p-values, and effect size (η^2) quantified whether monthly variations were statistically significant, supporting process scheduling and improvement decisions.

Delivery Time Comparison for Mouse Category:
H_0: There is no significant difference in delivery times across months (1–12).
H_1: There is a significant difference in delivery times across months (1–12).

6.2 Statistical Results

The one-way ANOVA results showed a p-value < 0.001, indicating statistically significant differences between months at the 0.05 significance level. The effect size ($\eta^2 = 0.0735$) represents a small-to-medium impact. The highest mean delivery time occurred in Month 12 (24.36 hours) and the lowest in Month 1 (18.99 hours). Although significant, this variation explains only a modest portion of the total variability, suggesting that other operational factors also contribute. Continuous monitoring and improved scheduling during high-demand periods could help stabilise performance.

```
               Df Sum Sq Mean Sq F value Pr(>F)    
Month           1  56925   56925    1631 <2e-16 ***
Residuals   20660 720924      35                   
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table 2: Results of Anova table printed in R

6.3 Interpretation and Conclusion

The ANOVA results for the Mouse product category indicate that while statistically significant month-to-month variations exist, the practical differences are relatively modest. The observed range of mean delivery times—from approximately 19 hours in January to 24 hours in December—suggests that the process generally remains stable, with slight delays occurring during peak operational months. These fluctuations may be attributed to seasonal workload increases, supplier scheduling, or end-of-year demand surges rather than underlying inefficiencies in the delivery system itself. The small-to-medium effect size further supports the conclusion that the delivery process is largely consistent across the year. Nonetheless, maintaining proactive monitoring, adjusting staffing levels, and reviewing logistics during higher-volume periods could help further minimise variability and sustain reliable performance across all months.

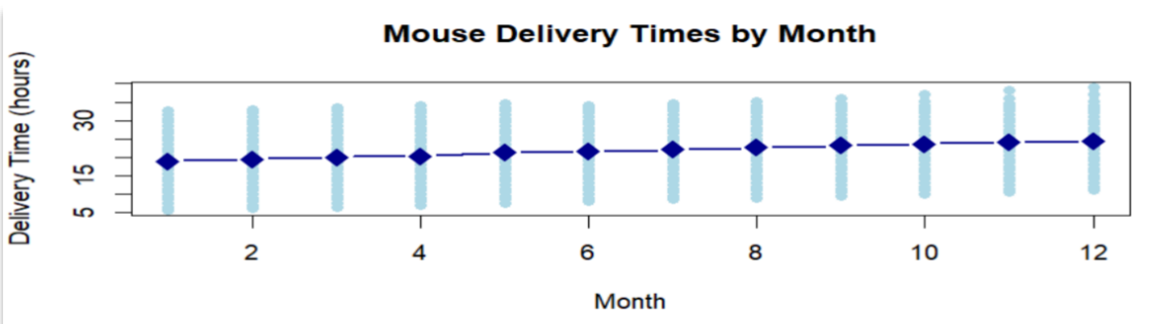

Figure 46:Mouse delivery times by Month

Part 7: Service Reliability and Staffing Optimisation

7.1 Methodology

If 15 or more workers are on duty, the service is considered reliable. To determine the total number of reliable days, the days with 15 and 16 workers present are summed. Based on the data, 96 days had 15 workers, and 270 days had 16 workers, resulting in 366 reliable days out of 397 observed. To estimate the annual expectation, the proportion of reliable days is calculated and multiplied by 365 days per year: (366 ÷ 397) × 365 = 336.5 reliable days per year. This corresponds to approximately 336–337 reliable service days annually, or about 28–29 days of potential service issues.

7.2 Staffing Profit Optimisation (Binomial Model)

7.2.1 Methodology

The information provided in the final report was used to model daily worker attendance and determine optimal staffing under uncertainty. Probability of presence (p = 0.974) was estimated from empirical data, and binomial modelling (dbinom(), pbinom()) simulated the likelihood of different attendance outcomes. Total monthly cost was modelled as, $\text{Total Cost} = \text{Salary Cost}(N) + \text{Loss Cost}(1–R(N)$,where R(N) represents reliability. Iterative computation identified the staffing level (N = 17) that minimized total expected cost, balancing additional wage expenditure (R25 000 per worker) against avoided service losses (R20 000 per unreliable day). Plots generated in ggplot2 depicted the cost curve and optimal point, providing a quantitative reliability-cost decision framework for managerial planning.

7.2.2 Analysis and Results

Worker attendance data reveals maximum observed presence of 16 workers, establishing the current staffing level. Daily attendance exhibits variation, with observations ranging from 12 to 16 workers present across 397 operating days. The weighted average attendance equals (12×1 + 13×5 + 14×25 + 15×96 + 16×270) / 397 = 15.58 workers, yielding individual attendance probability of p = 15.58/16 = 0.974.

Service reliability requires minimum 15 workers present. This threshold-based reliability criterion allows modelling of daily attendance as a binomial random variable X ~ Binomial (n, p=0.974), where X represents the number of workers present on any given day. The probability of service failure equals P(X < 15), representing insufficient staffing to maintain operational standards.

Total monthly cost comprises two components: fixed personnel cost and expected service failure penalties.

The cost function is expressed as:

C(n) = (n × R25,000) + (P(X < 15) × 30 days × R20,000)

For current staffing of n=16 workers, P(X < 15) = 0.0636, yielding monthly cost of C(16) = R400,000 + R38,179 = R438,179.

Increasing staffing to n=17 workers reduce failure probability to P(X < 15) = 0.0091, yielding C(17) = R425,000 + R5,443 = R430,443. This represents R7,736 monthly savings despite R25,000 additional salary cost, as the R32,736 reduction in expected failure costs exceeds the marginal labour expense. Further increases to n=18 workers would reduce failure costs by only R4,819, failing to offset the R25,000 salary increment.

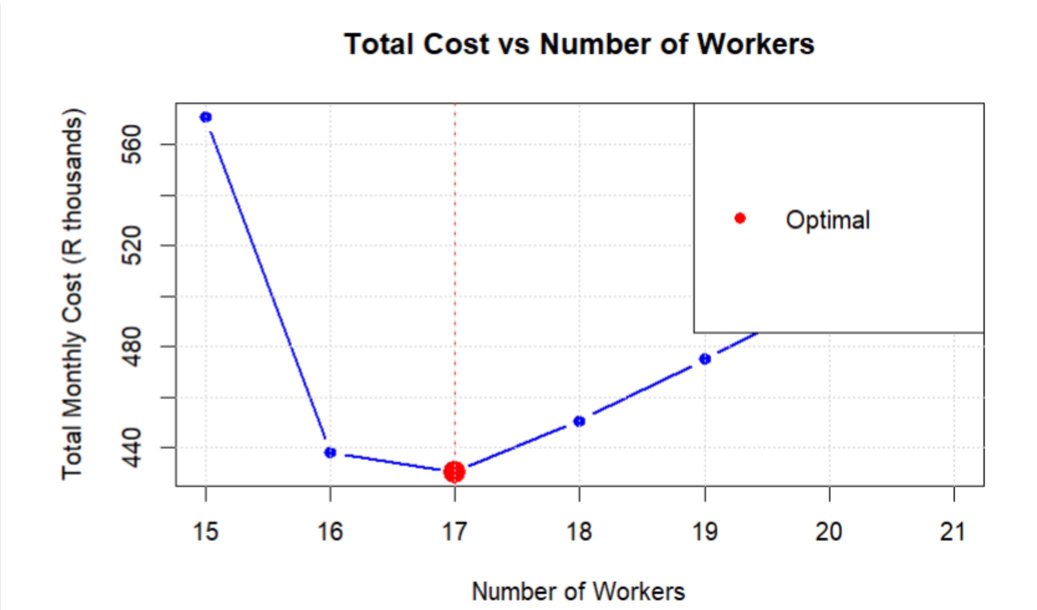

Figure 47: Total cost vs number of workers

Figure 47 illustrates the total cost function across staffing levels from 15 to 21 workers. The curve exhibits U-shaped morphology characteristic of optimization problems balancing competing cost components. At n=15 workers, total cost reaches R560,000 due to high service failure probability. Cost decreases sharply to R438,19 at n=16 workers as failure probability declines substantially. The minimum occurs at n=17 workers (R430,443), marked by the red optimal point. Beyond this minimum, costs increase linearly at R25,000 per additional worker, as failure probability approaches zero and fixed labour costs dominate the cost function.

The sharp cost reduction from 15 to 17 workers reflects the nonlinear relationship between staffing and reliability in binomial systems. Small staffing increases near the reliability threshold (15 workers) produce disproportionate reliability improvements. Beyond the optimal point, diminishing marginal returns to reliability render additional workers economically inefficient, as failure probability becomes negligibly small and cannot justify further labour expenditure.

Conclusion:

The analyses conducted across all sections provide an integrated understanding of quality control, process capability, and operational optimization. The SPC results highlight areas for process improvement, while the staffing optimization confirms the strong link between workforce levels and service profitability. Together, these findings demonstrate the application of quantitative decision-making in achieving operational excellence, fulfilling the ECSA GA4 requirement for engineering analysis and problem-solving competence.

References

- Engineering Council of South Africa (ECSA). 2025. *QA344 2025 ECSA GA4 Outcome in Industrial Engineering: Project Instructions and Assessment Rubric.* Stellenbosch: Department of Industrial Engineering, Stellenbosch University.
- Stellenbosch University. 2025. *QA344 Statistics: Statistical Methods in Quality Assurance.* Stellenbosch: Department of Industrial Engineering, Stellenbosch University.
- STHDA. n.d. *MANOVA Test in R: Multivariate Analysis of Variance.* [online] Available at: http://www.sthda.com/english/wiki/manova-test-in-r-multivariate-analysis-of-variance [Accessed 22 October 2025].
- R Core Team. 2024. *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing. Available at: https://www.R-project.org [Accessed 22 October 2025].
- YouTube. n.d. *Introduction to Statistical Process Control (SPC).* [video online] Available at: https://www.youtube.com/watch?v=0GfBSuwhUwI [Accessed 22 October 2025].