ECSA

24938432
Morné Stephan Burger

TABLE OF CONTENTS

INTRODUCTION

This report, completed as part of the QA344 module, demonstrates achievement of Graduate Attribute 4 (GA4) of the Engineering Council of South Africa (ECSA). GA4 focuses on the ability to identify, analyse, and solve complex engineering problems using data-driven approaches. The project applies statistical reasoning, coding, and engineering judgement to evaluate process control, operational performance, and service reliability. The analysis begins with data preparation and descriptive statistics, followed by Statistical Process Control (SPC) using $\bar{X}$–S charts to assess process stability and capability. The likelihood of Type I (manufacturer's) and Type II (consumer's) errors is then calculated to evaluate the reliability of control decisions. Design of Experiments (DOE) and Analysis of Variance (ANOVA) are used to compare delivery performance across years and months, identifying any significant operational differences. Finally, an optimisation model is developed to balance efficiency and profitability, while a binomial reliability model estimates service reliability and determines optimal staffing levels. Overall, the project highlights how quantitative analysis and statistical tools can support evidence-based decision-making, improving process capability, reliability, and profitability within an engineering environment.

DATA OVERVIEW

Customer_data has 5000 observations with 5 variables: CustomerID, Gender, Age, Income and City. Products_data has 60 observations with 5 variables: ProductID, Category, Description, SellingPrice and Markup. Products_Headoffice has 360 observations with 5 variables: ProductID, Category, Description, SellingPrice and Markup. There are two instances where there are duplicate product IDs with different descriptions, duplicate IDs can cause serious inaccuracies in the sales data and the final data analysis.

DATA INACCURACY

During the initial review of the *products_data.csv* and *products_Headoffice.csv* files, several inconsistencies and structural errors were identified that affected the accuracy of profit calculations and category summaries. The following key inaccuracies were detected:

Products from line 11 onwards in each category contained incorrect prefixes such as **"**NA011**"**, **"**NA012**"**, etc. These should have matched their true product type (SOF for *Software*, KEY for *Keyboard*, MOU for *Mouse*). As a result, many products were categorised under the wrong type, which caused average profit and markup values to be allocated incorrectly across categories. Only the first ten products of each type contained correct selling price and markup values. From product 11 onwards, incorrect price and markup values were repeated or mixed between types (for example, *Software* items showing *Monitor* pricing). This led to inflated and inconsistent profit calculations — especially within the *Software*, *Keyboard*, and *Mouse* categories. In the Head Office file, the same price–markup combination appeared across all product types, creating uniform but unrealistic profits (around R700–R800 per item). This pattern ignored product cost differences and masked the true variation in profitability between categories. Several products' Category entries did not correspond to their ProductID prefix (e.g., "SOF009" recorded under "Laptop"). This misalignment caused aggregate summaries (such as "Average Profit per Category") to display distorted results and false category averages.

Table 1:Old products_data vs New products_data2025

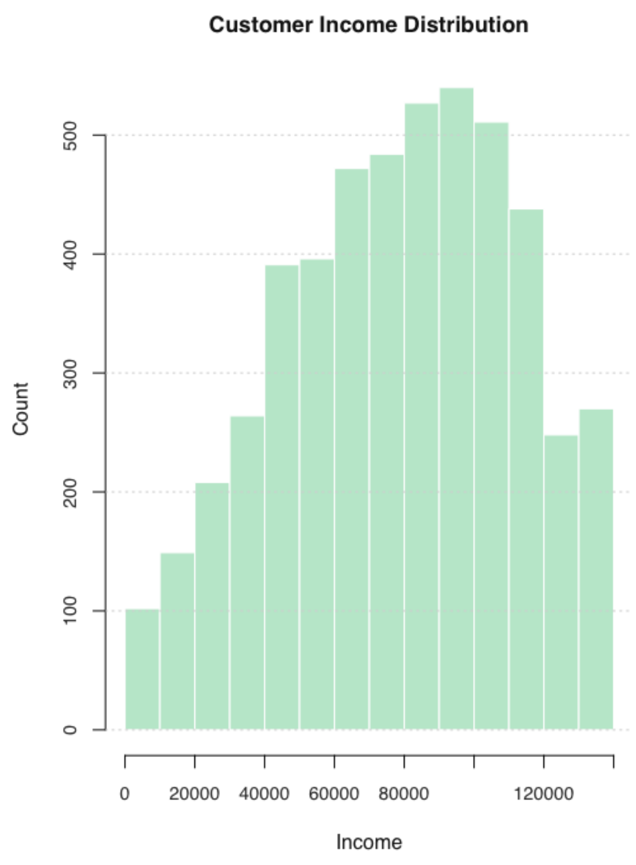
CATEGORY	OLD DATA – AVG PROFIT (R)	NEW DATA – AVG PROFIT (R)	CHANGE (R)	% CHANGE
LAPTOP	937.32	2 763.62	▲ 1 826.30	+195 %
MONITOR	837.63	1 193.05	▲ 355.43	+42 %
CLOUD SUBSCRIPTION	553.76	167.57	▼ 386.18	−70 %
KEYBOARD	725.26	124.08	▼ 601.18	−83 %
SOFTWARE	489.08	68.42	▼ 420.66	−86 %
MOUSE	840.04	66.34	▼ 773.70	−92 %

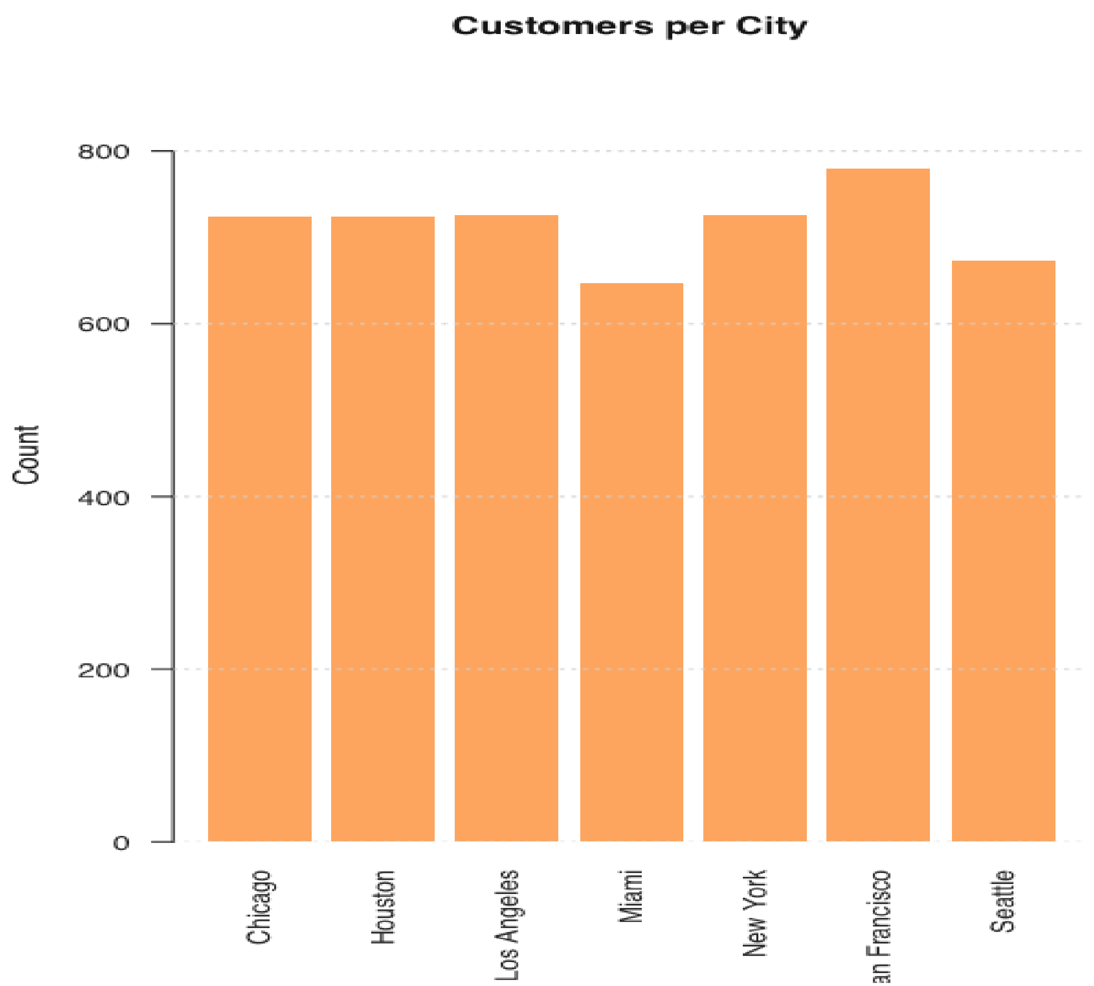
Table 2: Old products_Headoffice VS New products_Headoffice2025

CATEGORY	OLD DATA – AVG PROFIT (R)	NEW DATA – AVG PROFIT (R)	CHANGE (R)	% CHANGE
CLOUD SUBSCRIPTION	790.98	156.57	▼ 634.41	−80 %
KEYBOARD	684.30	156.42	▼ 527.88	−77 %
LAPTOP	755.49	156.11	▼ 599.37	−79 %
MONITOR	701.52	154.26	▼ 547.26	−78 %
MOUSE	763.24	154.99	▼ 608.25	−80 %
SOFTWARE	788.85	154.58	▼ 634.27	−80 %

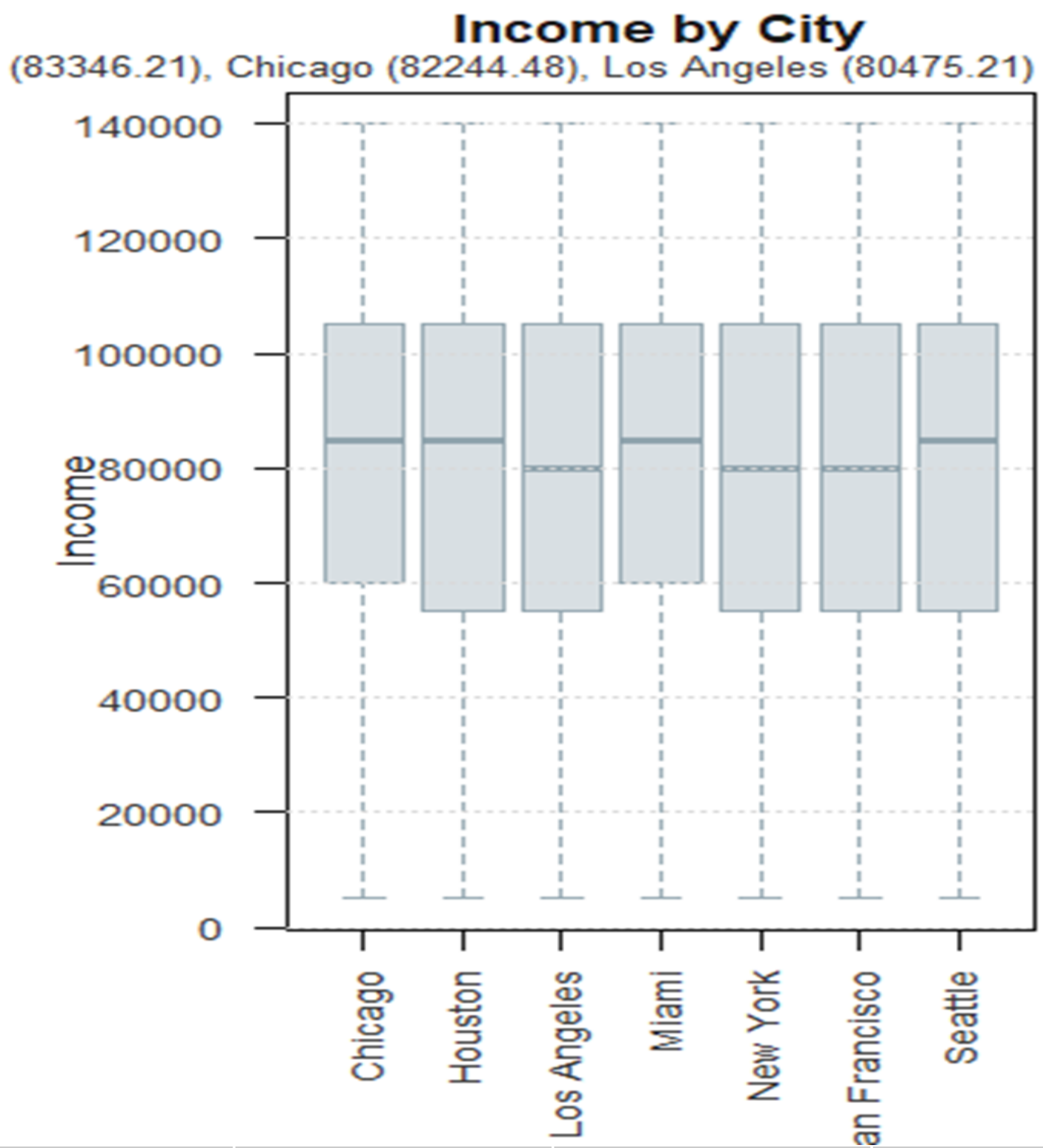
PRODUCTS_DATA

- Income is moderately skewed; key band clusters around the median with a right tail. ***Median* ≈ 85,000**; ***mean* ≈ 80,797**; ***IQR* 55,000–105,000**.

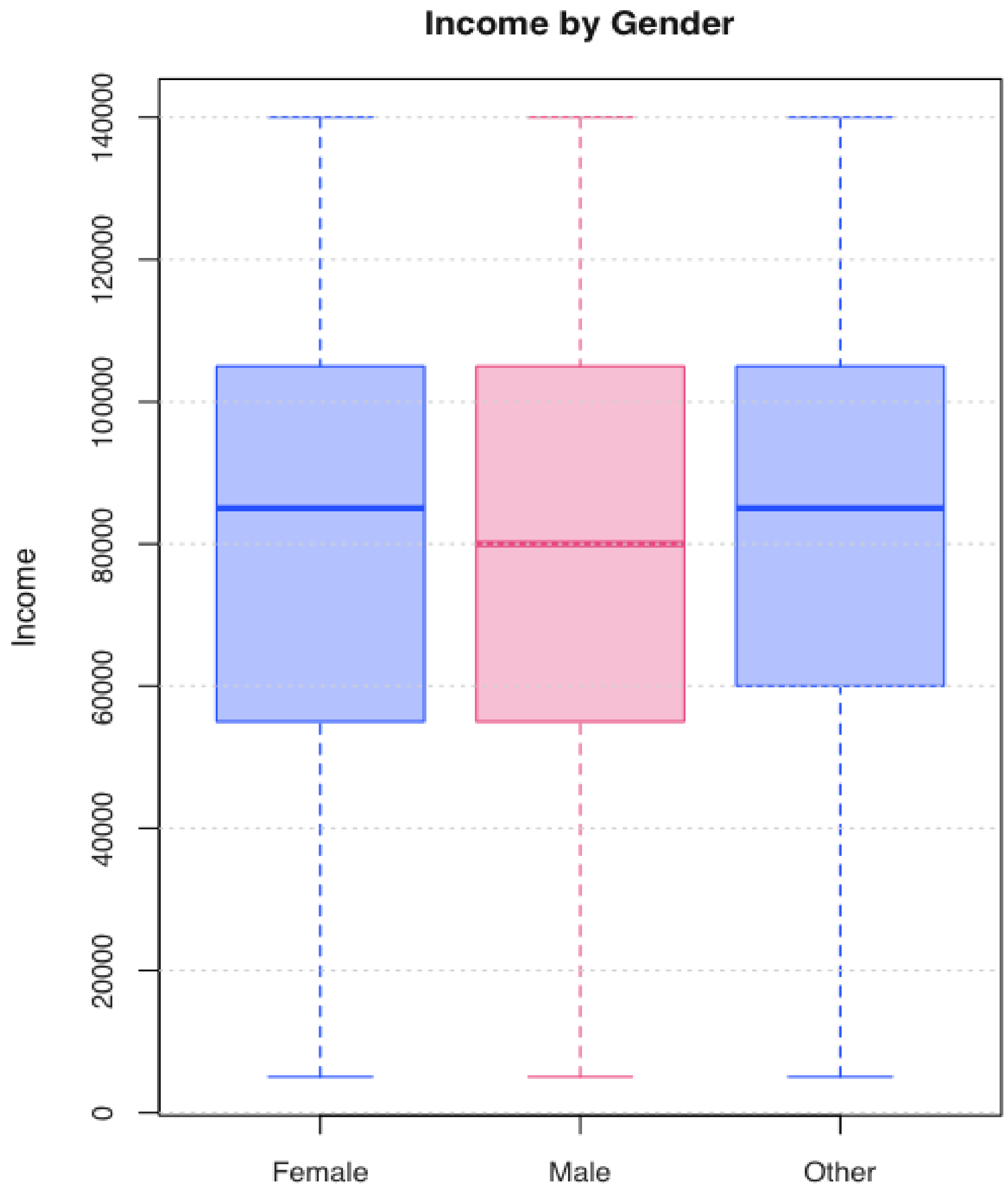




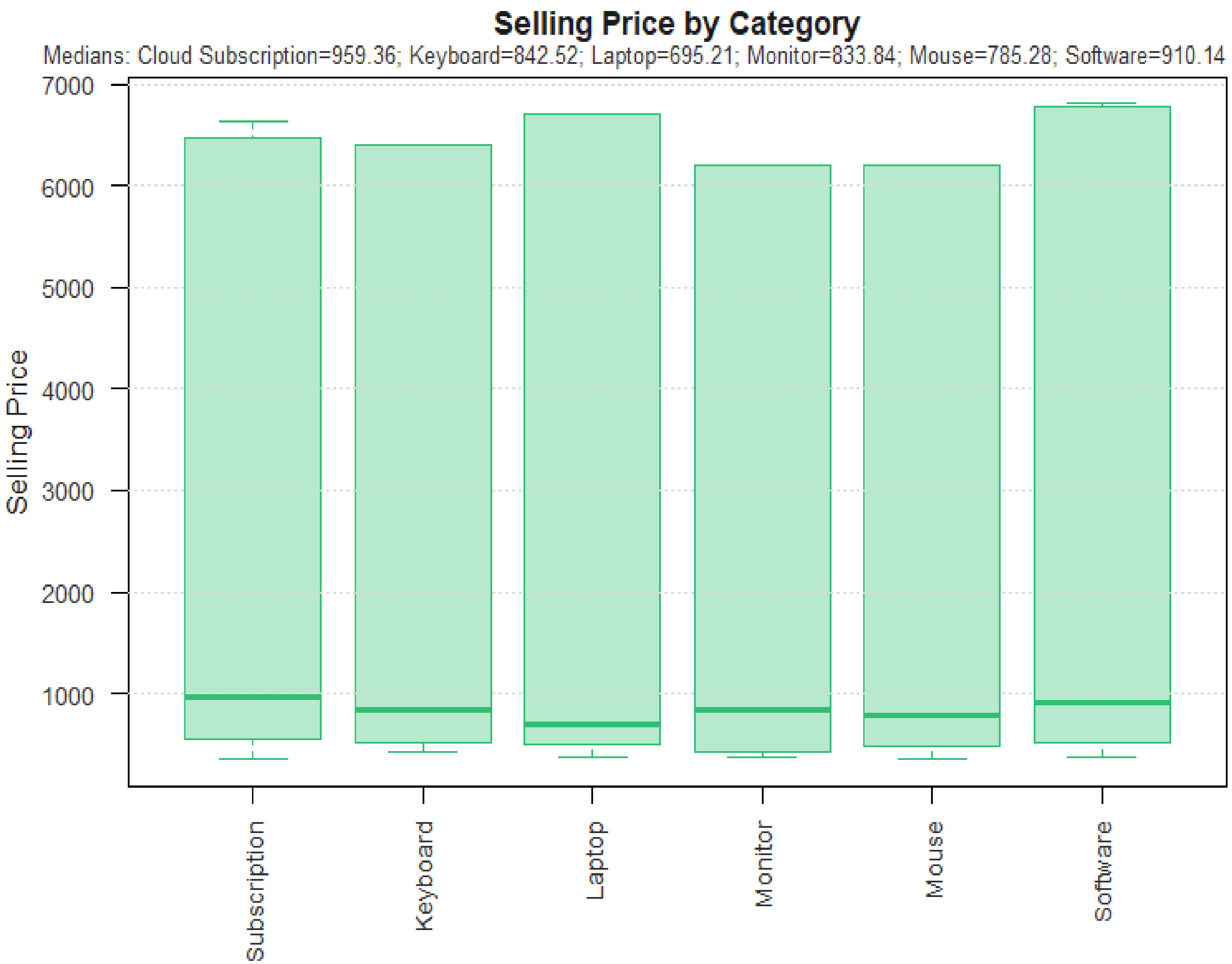
City	San Francisco	Los Angeles	New York	Chicago	Houston	Seattle	Miami
Total Customers per city	780	726	726	724	724	673	647



City	San Francisco	Los Angeles	New York	Chicago	Houston	Seattle	Miami
Total Income per city	62285000	58425000	57900000	59545000	58100000	53805000	53925000
Average Income per city	79852.56	80475.21	79752.07	82244.48	80248.62	79947.99	83346.21



Gender	Male	Female	Other
Average Income	80870,21	80816.20	80871.56
Total Customers	2350(47%)	2432(48,6%)	218(4,4%)



Category	Cloud Subscription	Keyboard	Laptop	Monitor	Mouse	Software
Average selling price	3691.861	4638.172	5217.545	5014.170	4585.465	3814.344

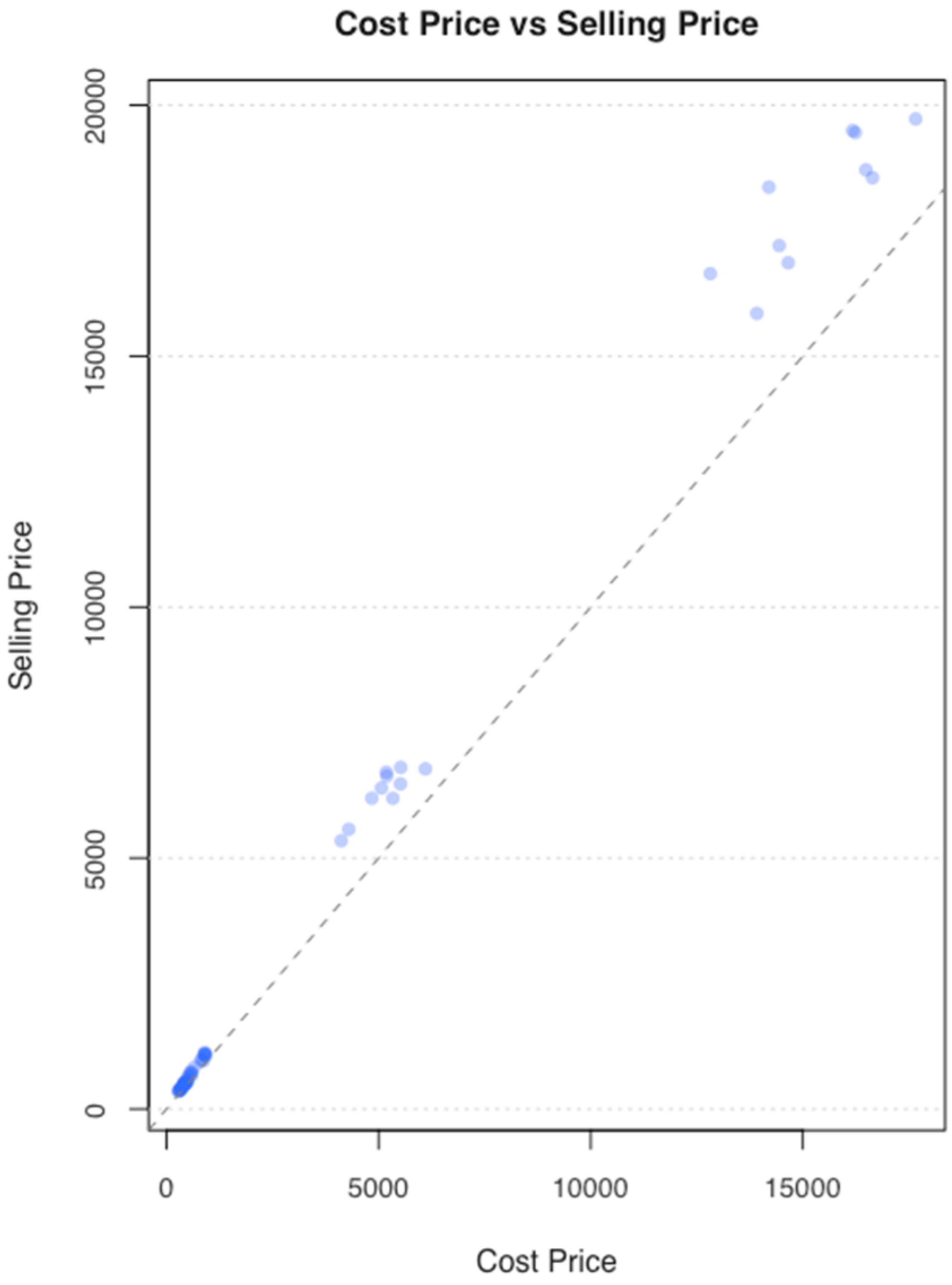
5 LEAST PROFITABLE PRODUCTS

9	SOF009	Laptop	540.41	55.04086
57	MOU057	Laptop	394.30	53.82854
5	SOF005	Keyboard	516.15	51.19189
2	SOF002	Cloud Subscription	505.26	47.72129
54	MOU054	Mouse	417.40	44.32170

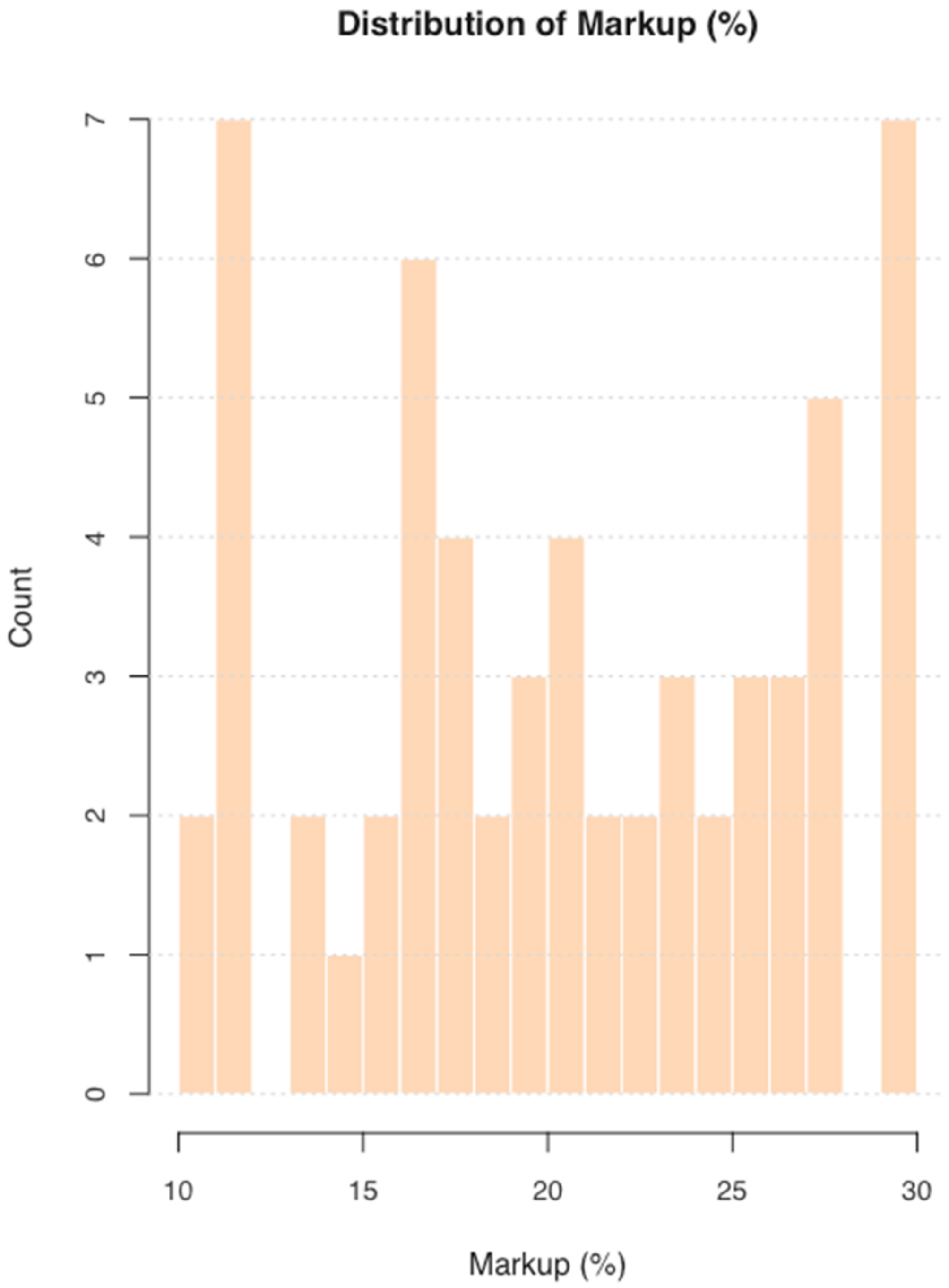
TOP 5 MOST PROFITABLE PRODUCTS

24	LAP024	Mouse	18366.92	4167.523
22	LAP022	Monitor	16644.21	3825.194
21	LAP021	Laptop	19494.91	3321.930
23	LAP023	Keyboard	19452.72	3215.057
27	LAP027	Laptop	17202.28	2759.933

Category	Cloud Subscription	Keyboard	Laptop	Monitor	Mouse	Software
Average profit	553.7552	725.2574	937.3168	837.6254	840.0377	489.0840
Average Cost price	3138.106	3912.915	4280.228	4176.545	3745.427	3325.260
Average Markup	20.553	20.161	20.623	20.727	20.668	20.038

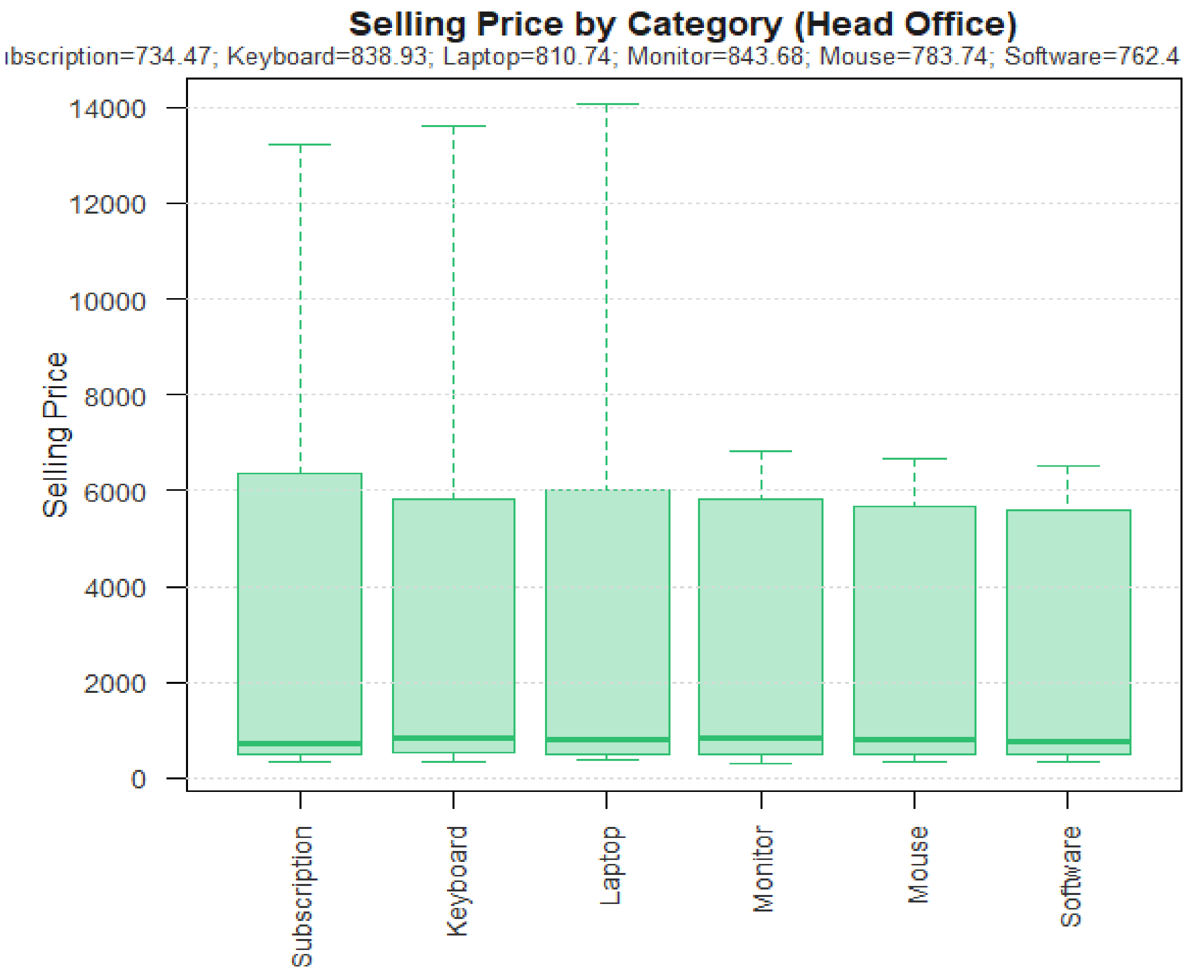


("Cost Price vs Selling Price") shows a clear positive relationship between cost and selling price, with all points lying above the dashed equality line—indicating that selling prices consistently exceed cost prices. The data clusters into distinct groups, suggesting product categories with different price ranges, and the consistent gap above the line reflects stable markups across these categories.



The histogram shows a wide and uneven distribution of markup percentages ranging from 10% to 30%, with distinct peaks around 10%, 17%, and 30%. This pattern suggests a bimodal or segmented pricing strategy, where products tend to cluster either at lower markups (likely cost-sensitive or high-volume items) or at higher markups (possibly premium or niche products), with fewer items priced in the mid-range. Overall, the data indicates varied pricing behavior across products, possibly reflecting differences in market positioning, product type, or customer segment focus.

PRODUCTS_HEADOFFICE



Category	Cloud Subscription	Keyboard	Laptop	Monitor	Mouse	Software
Average selling price	4386.710	4380.485	4305.739	4456.745	4478.900	4457.193

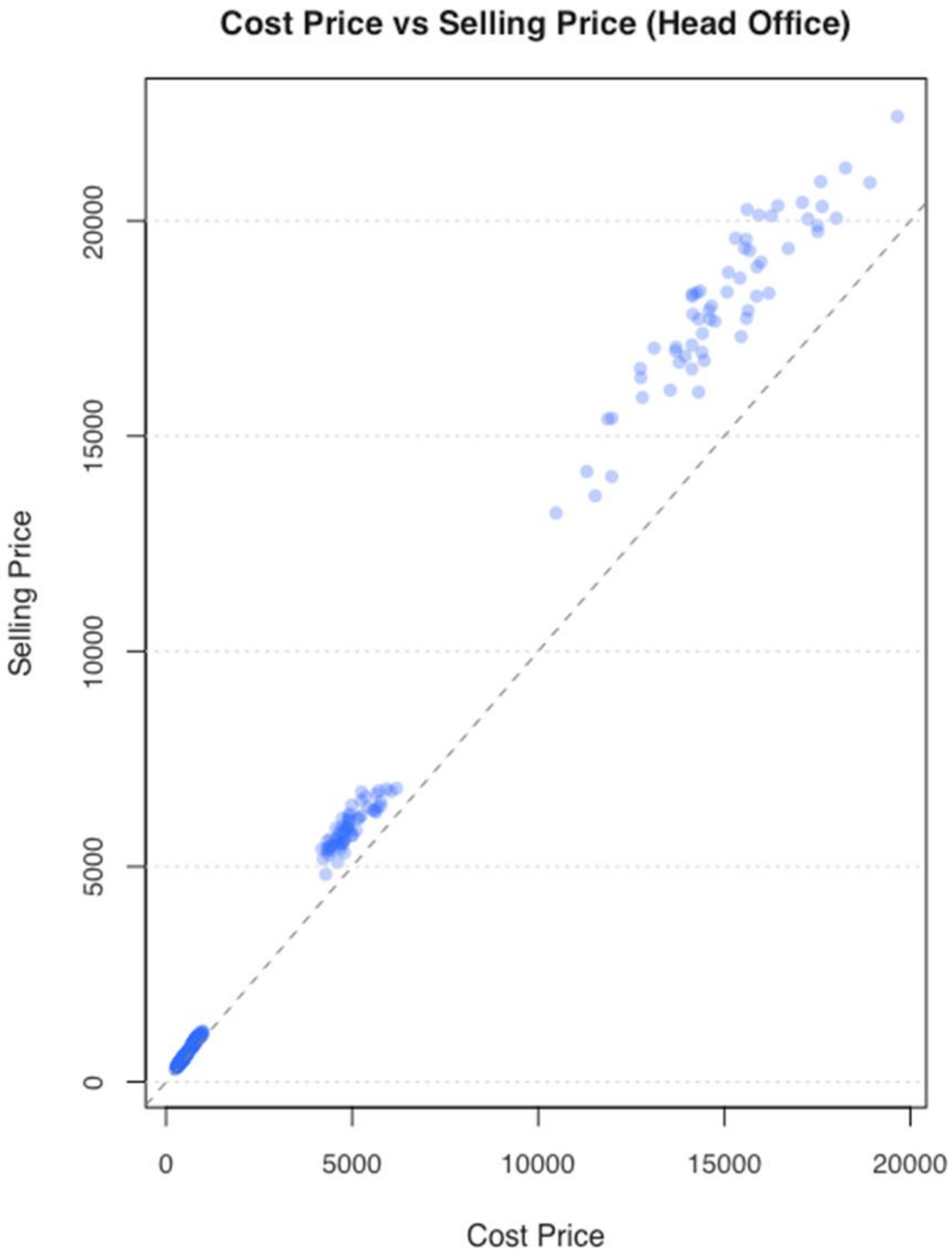
Top 5 most profitable products

22	NA022	Software	4648.070
322	NA022	Mouse	4286.644
267	NA027	Keyboard	4197.819
90	NA030	Cloud Subscription	4159.754
26	NA026	Software	4103.496

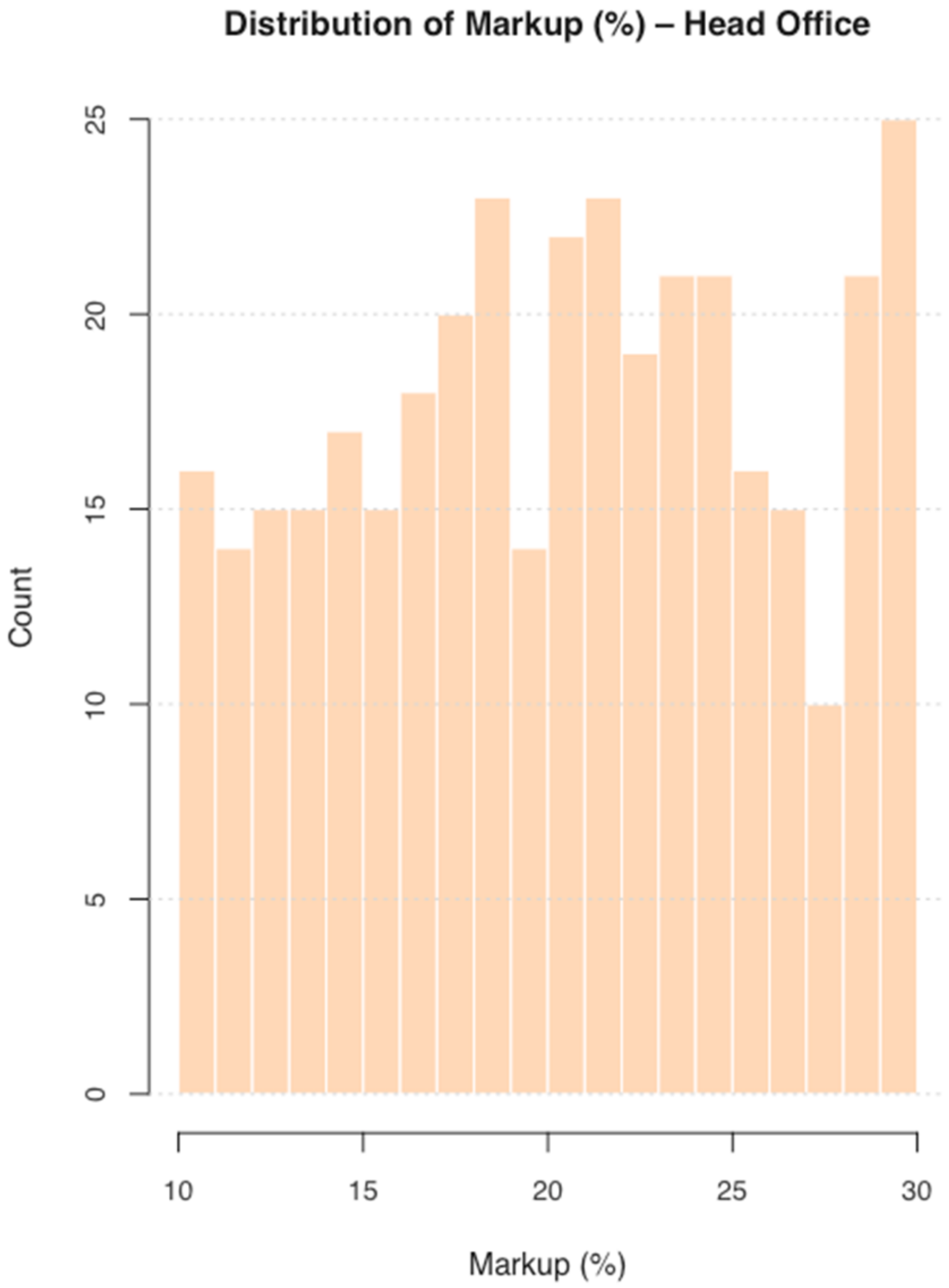
5 Least profitable products

115	NA055	Cloud Subscription	39.79854
237	NA057	Monitor	39.68264
351	NA051	Mouse	39.56545
178	NA058	Laptop	37.29962
174	NA054	Laptop	36.30579

Category	Cloud Subscription	Keyboard	Laptop	Monitor	Mouse	Software
Average profit	790.9756	684.3018	755.4858	701.5204	763.2430	788.8482
Average Cost price	3595.734	3696.184	3550.253	3755.224	3715.657	3668.344
Average Markup	21.50000	19.95417	20.47517	19.44250	20.17967	20.76150



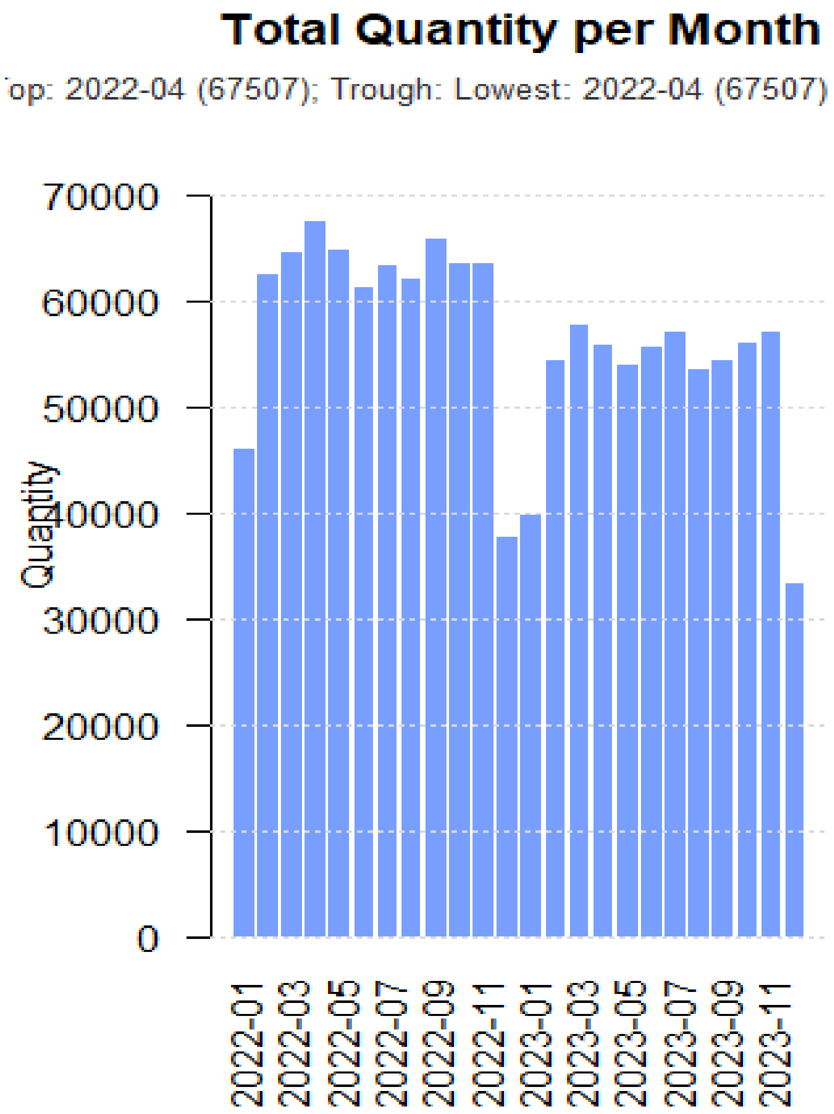
("Cost Price vs Selling Price – Head Office") displays a trend but with a larger sample size and tighter clustering. Most data points remain above the equality line, confirming consistent profitability. However, the denser spread at higher cost and selling price ranges (around 15,000–20,000) implies that the Head Office handles a greater volume of high-value transactions, maintaining steady markups while scaling across a wider product base.



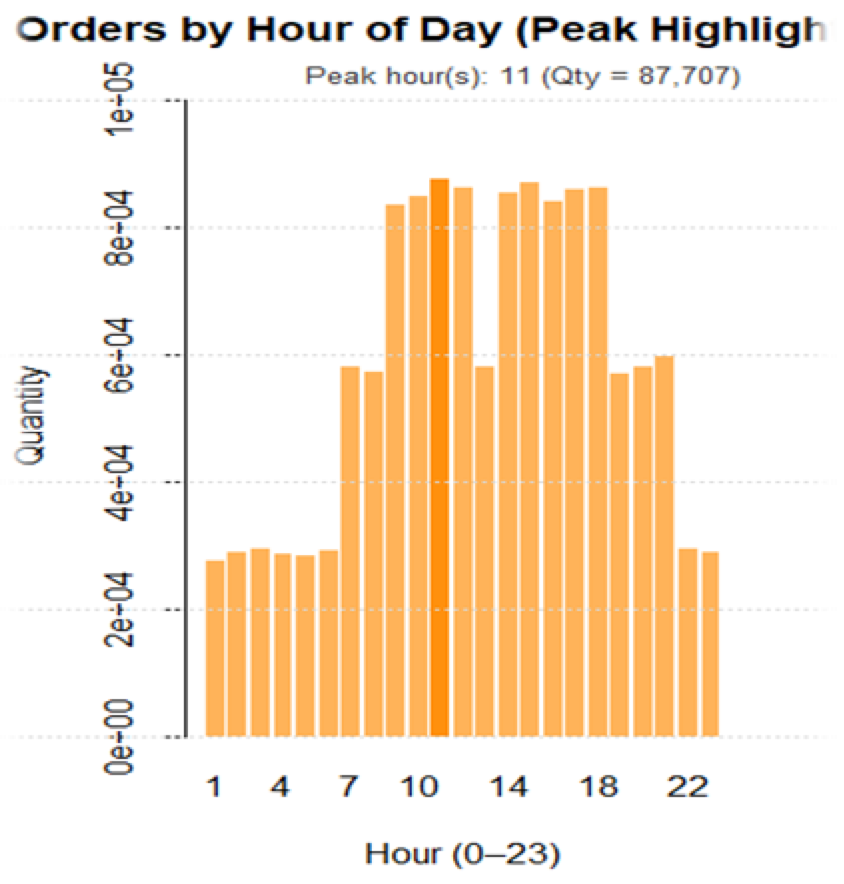
The histogram shows the distribution of markup percentages for the Head Office, ranging from 10% to 30%. The data appears relatively balanced but slightly skewed toward higher markups, with frequencies increasing steadily from around 15% upward and peaking near 30%. This suggests that while lower markups are still used, the majority of products or transactions tend to cluster around mid- to high-markup levels (20–30%), indicating a stronger pricing emphasis on higher-margin items. Overall, the pattern reflects a consistent yet upward-trending markup strategy across the Head Office dataset.

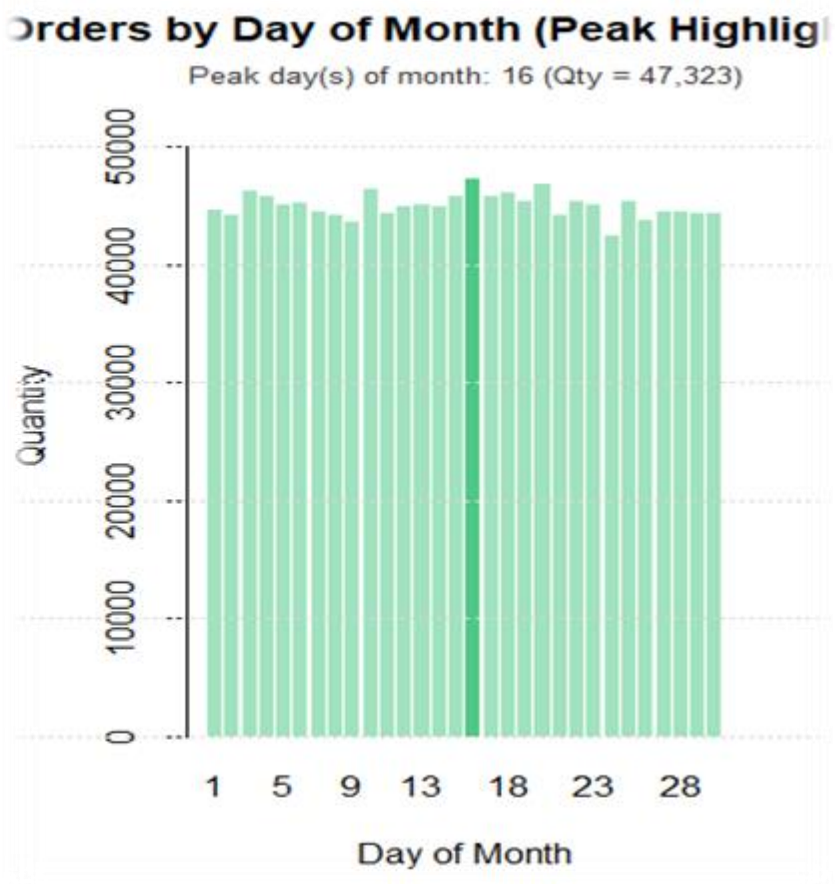
SALES2022AND2023

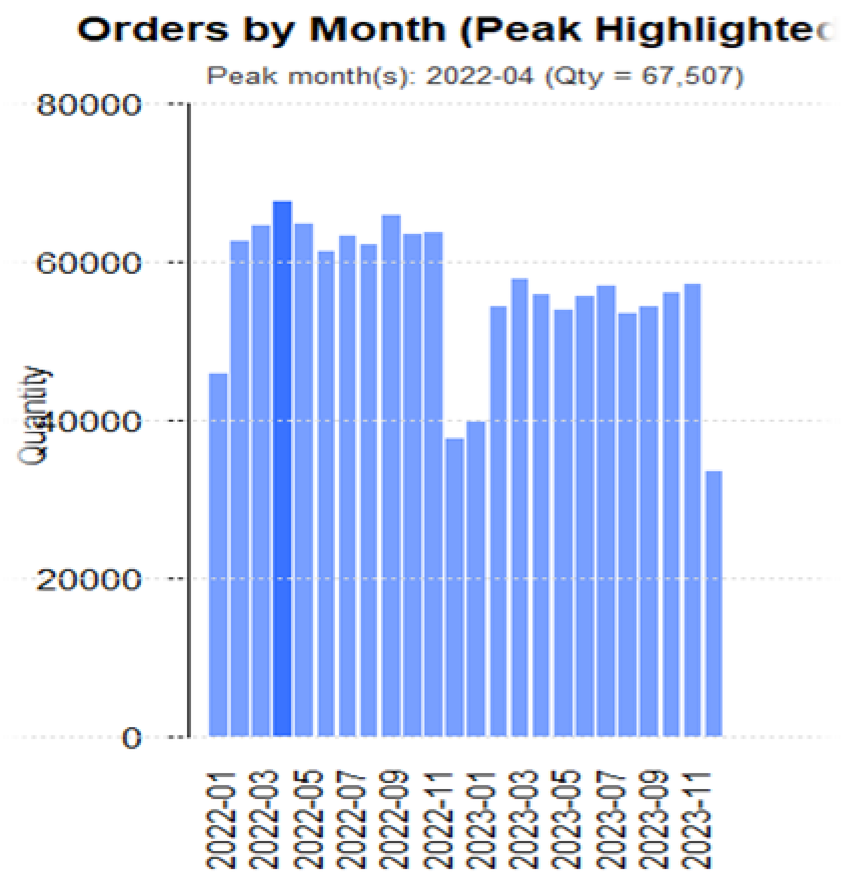
Year	Month	Quantity
2022	1	45912
2022	2	62551
2022	3	64471
2022	4	67507
2022	5	64743
2022	6	61243
2022	7	63206
2022	8	62084
2022	9	65828
2022	10	63443
2022	11	63544
2022	12	37609
2023	1	39771
2023	2	54248
2023	3	57680
2023	4	55826
2023	5	53915
2023	6	55557
2023	7	57014
2023	8	53486
2023	9	54403
2023	10	55949
2023	11	57025
2023	12	33332



The bar chart shows the **total quantity of sales per month** from January 2022 to November 2023. Overall, the data reveals consistently high sales volumes, with quantities generally ranging between **50 000 and 70 000 units per month**. A noticeable **peak occurred in April 2022** (around 67 507 units), while a **dip followed later in 2022**, before quantities stabilized again through 2023. The pattern suggests **steady operational activity** with short-term fluctuations, possibly reflecting seasonal demand changes or supply constraints. Despite the mid-period slowdown, the business maintained a **strong recovery trend** into 2023, demonstrating stable and resilient sales performance over the two-year period.







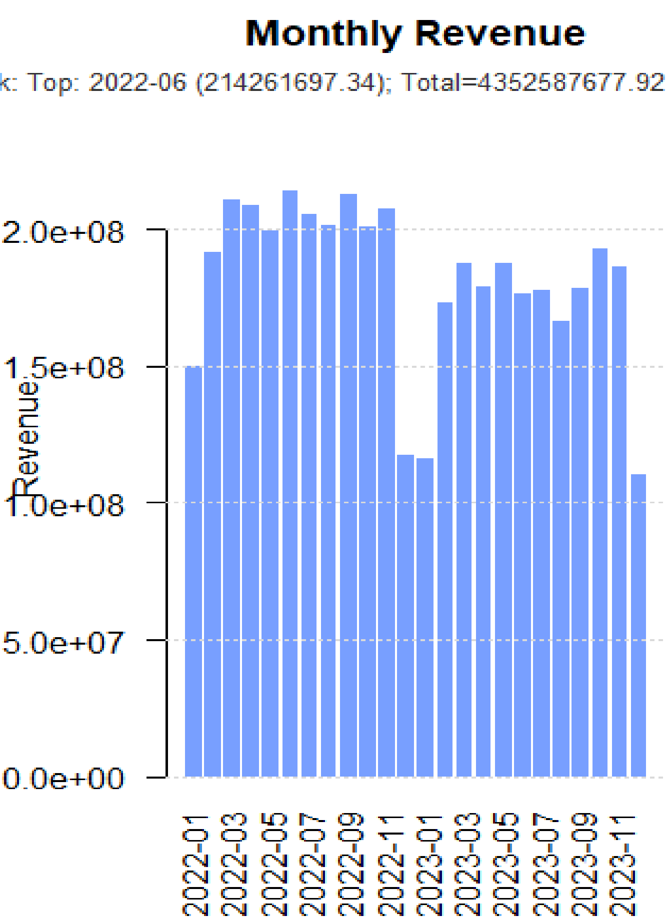
By Hour of Day: Orders rise sharply in the morning, peaking around **11:00**, when the quantity reaches approximately **87 700 units**.

By Day of Month: The distribution is relatively even, but the 16th day of the month stands out with a peak of about 47 300 units.

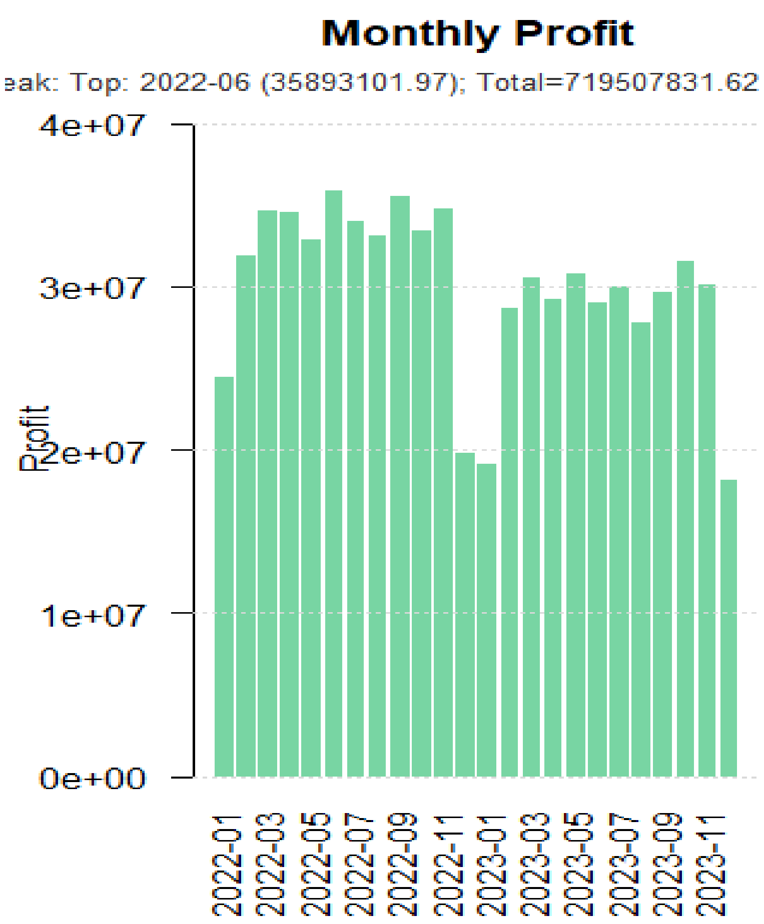
By Month: The **overall monthly trend** shows April 2022 as the highest month, with roughly **67 500 units sold**, followed by consistently strong months thereafter.

Top 10 Products by Quantity

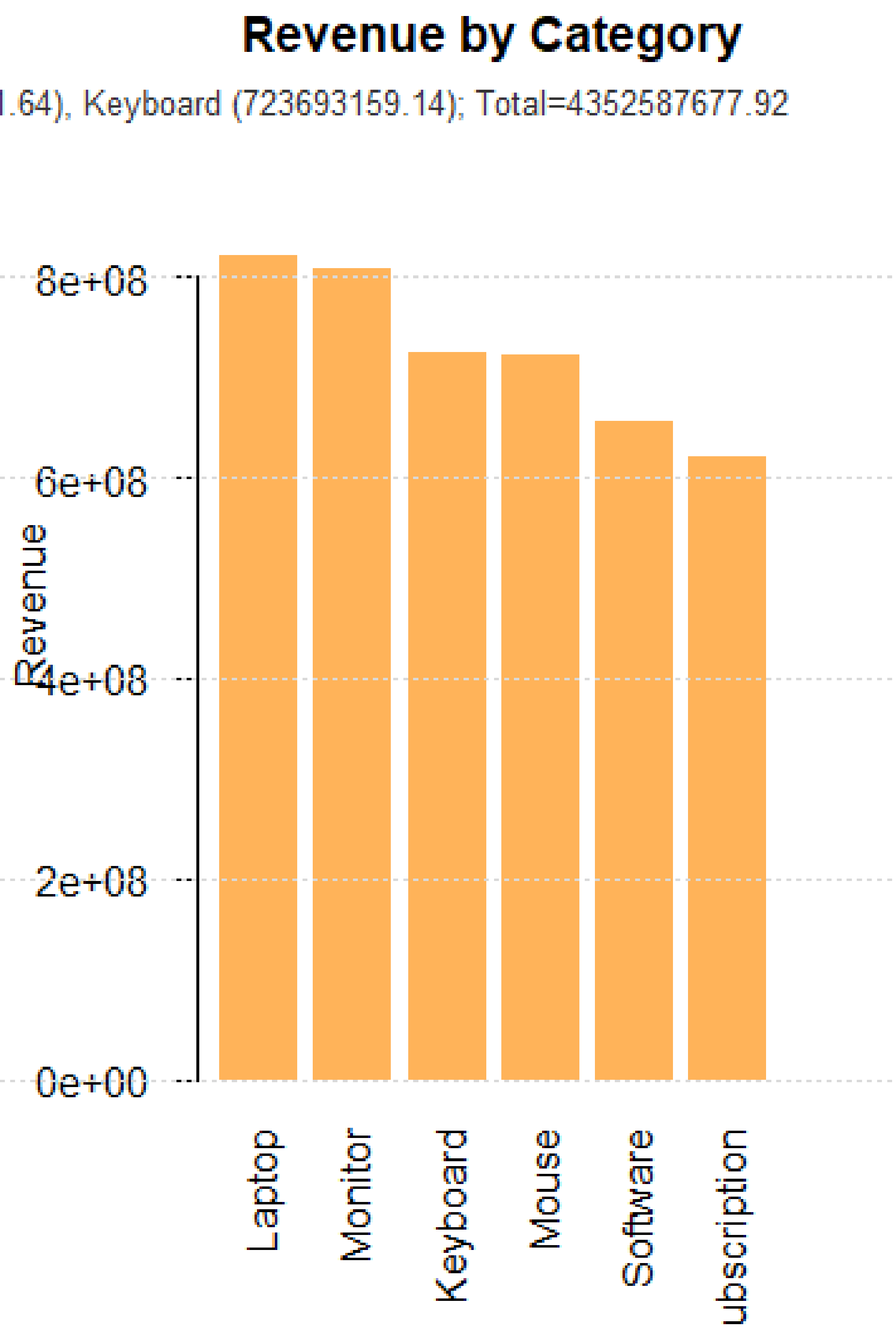
ProductID	Quantity
MOU059	29675
SOF001	29336
SOF004	29219
SOF010	29168
MOU058	28924
MOU054	28875
MOU052	28804
SOF007	28517
MOU057	28423
SOF005	28412



Year	Month	Revenue	Profit
2022	1	149734393	24448024
2022	2	191694994	31886100
2022	3	210876809	34693396
2022	4	208723699	34525374
2022	5	199709671	32865280
2022	6	214261697	35893102
2022	7	205320954	33994168
2022	8	201506307	33073181
2022	9	212735572	35514127
2022	10	200846511	33378920
2022	11	207365657	34764101
2022	12	117633753	19796099
2023	1	116495955	19121895
2023	2	173176846	28705111
2023	3	187470850	30602724
2023	4	179326944	29270273
2023	5	187405959	30798655
2023	6	176422858	28970578
2023	7	177727957	29949980
2023	8	166659156	27771247
2023	9	178074787	29635795
2023	10	193088672	31593534
2023	11	186181380	30092882
2023	12	110146296	18163287



After the mid-2022 peak, both revenue and profit experienced a **dip** before recovering through 2023, though not quite reaching earlier highs. This mid-period decline may reflect seasonal factors, supply constraints, or demand slowdowns. Overall, the graphs demonstrate **strong top-line performance**, healthy profitability, and consistent alignment between sales volume and earnings—indicating that operational efficiency and cost control remained steady throughout most of the two-year period.

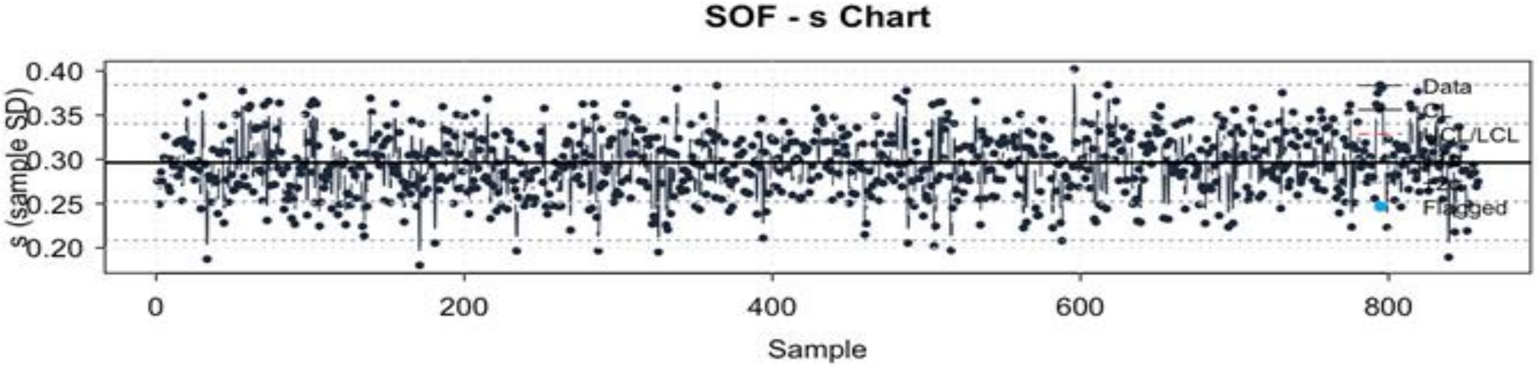


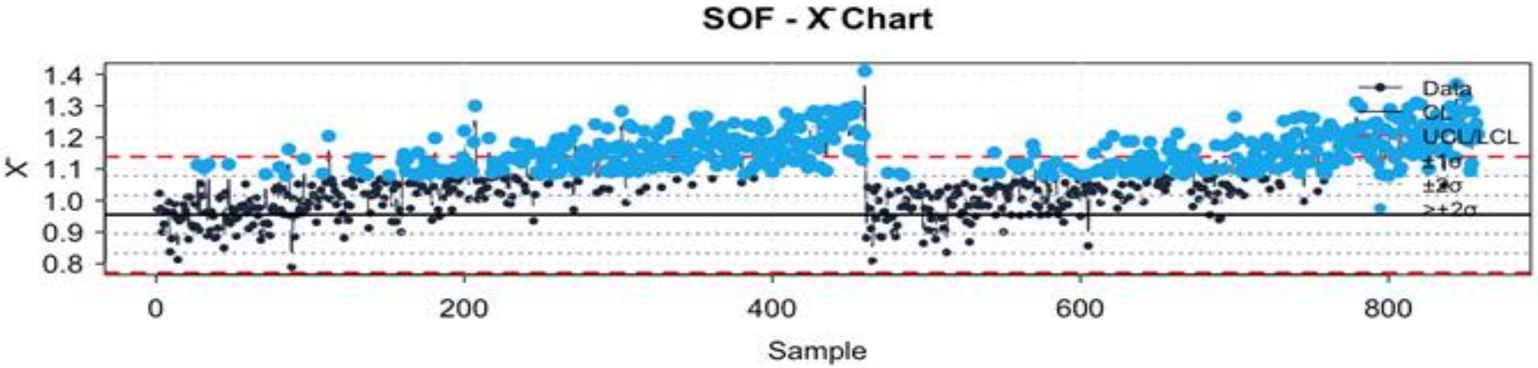
Category	Revenue	Profit
Laptop	821533851	149402796
Monitor	809104952	137086280
Keyboard	723693159	132148412
Mouse	721090260	116175960
Software	655365933	97305692
Cloud Subscription	621799523	87388692

SPC ANALYSIS

This section presents the $\bar{X}$–S control charts and key descriptive statistics for six products: SOF, CLO, KEY, MOU, MON, and LAP. Each figure is followed by a concise table summarizing process variation, control limits, and capability metrics, as well as a short interpretation of the observed stability and trends.

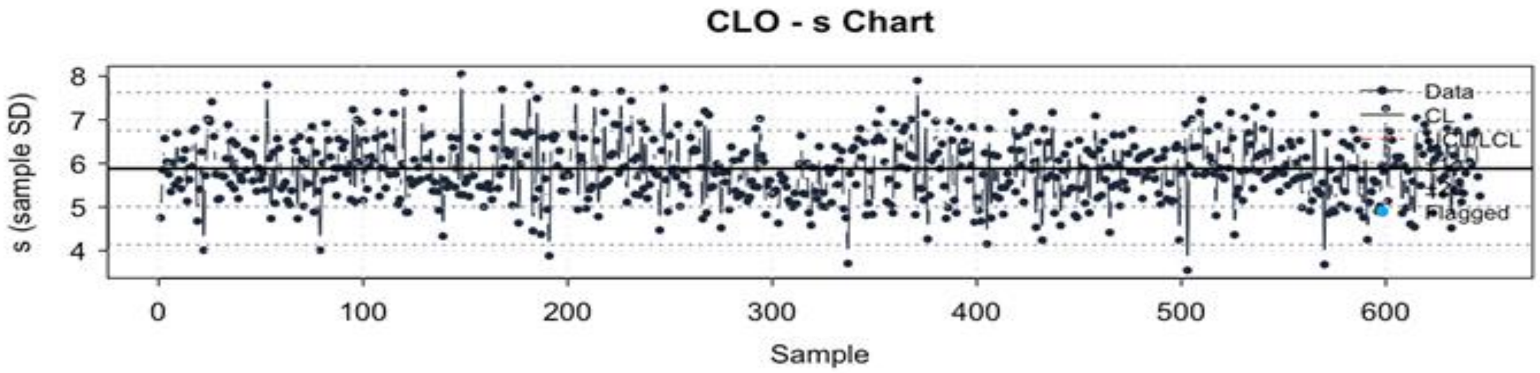
SOF

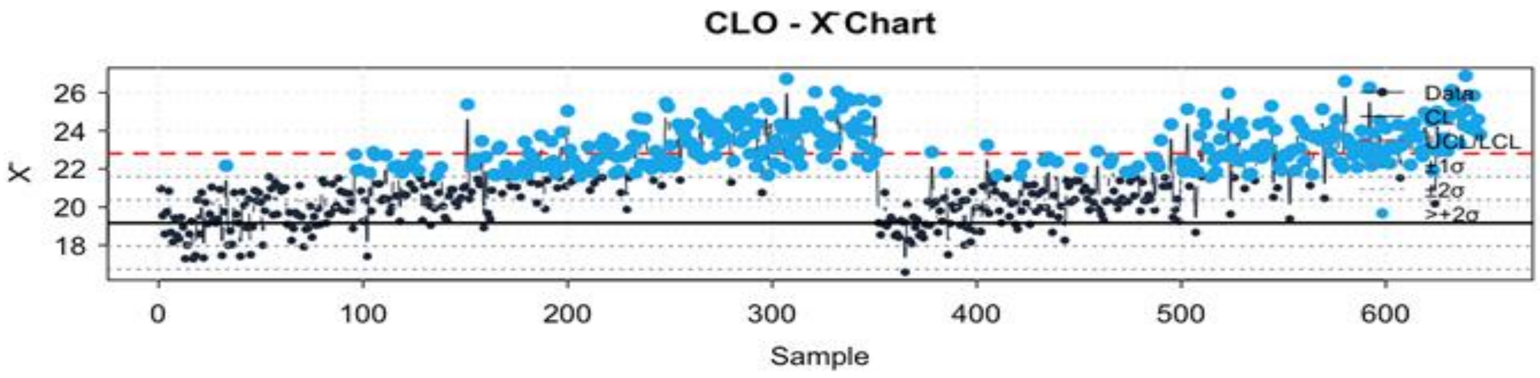




product	n_total_values	n_complete_samples	n_init_samples	xbar_CL	sbar	xbar_UCL	xbar_LCL
SOF	20613	858	30	0.955359722222222	0.296405596798734	1.13883478664064	0.771884657803806

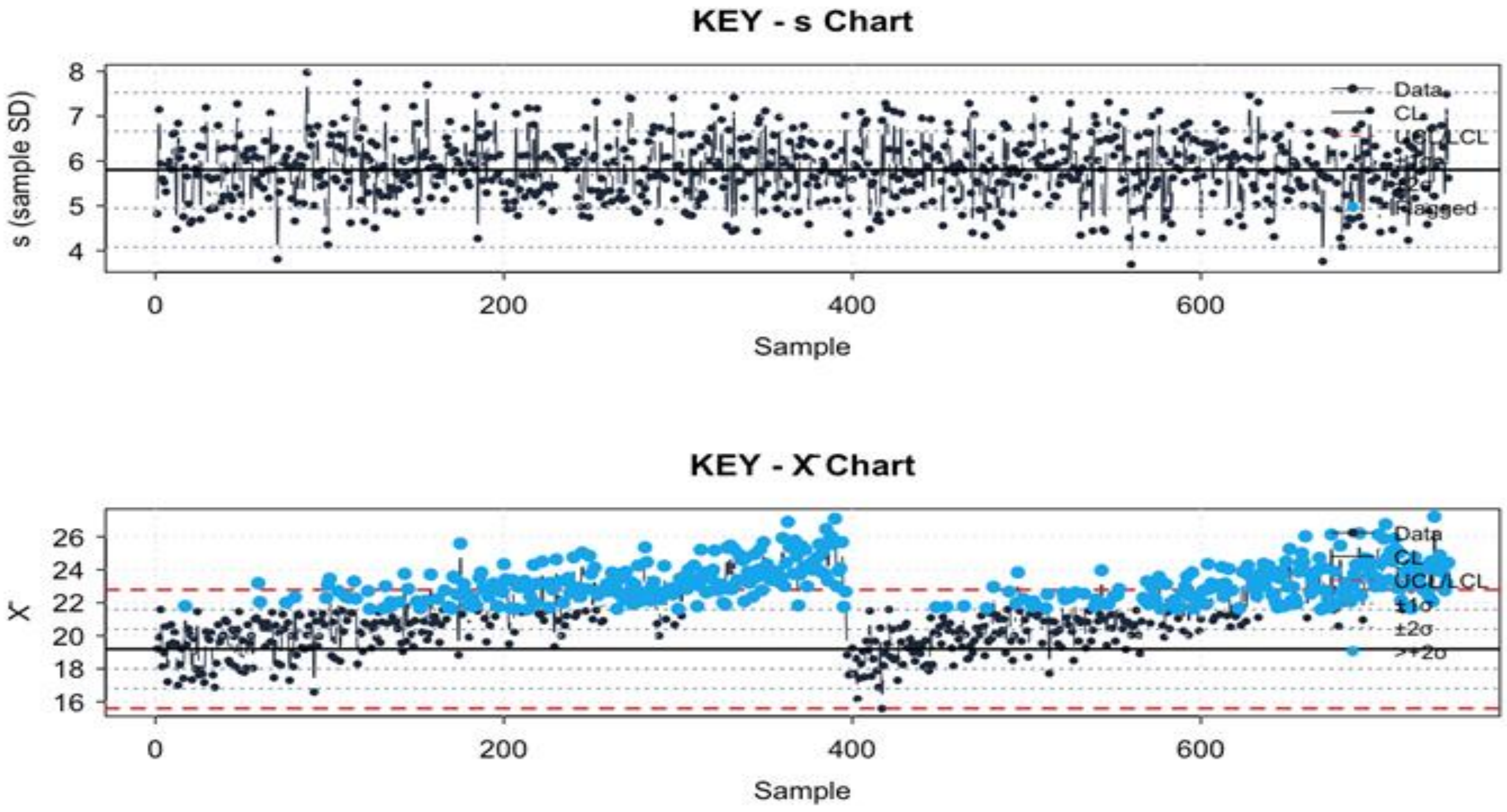
CLO





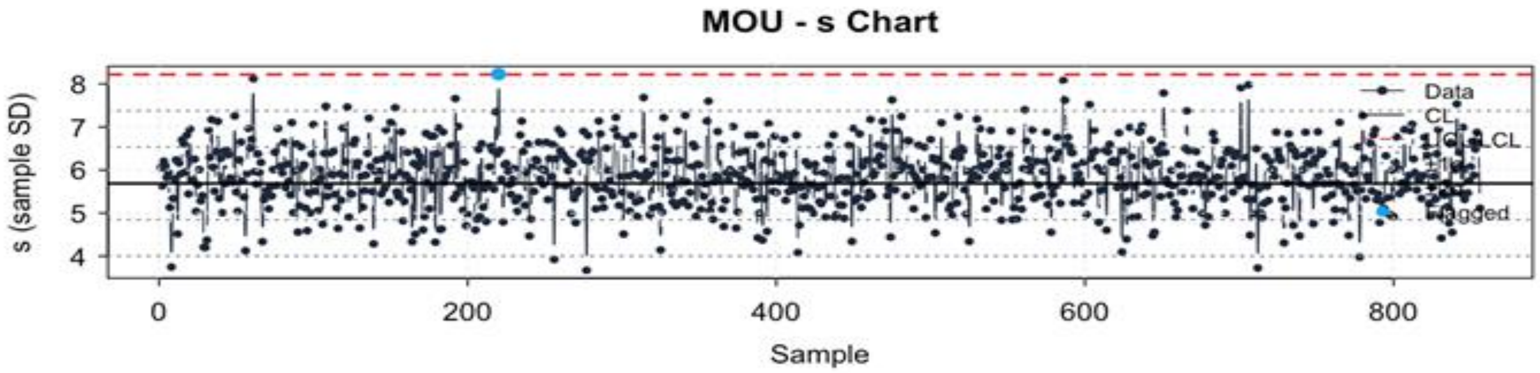
product	n_total_values	n_complete_samples	n_init_samples	xbar_CL	sbar	xbar_UCL	xbar_LCL
CLO	15519	646	30	19.1662222222222	5.88219731093651	22.8073023576919	15.5251420867525

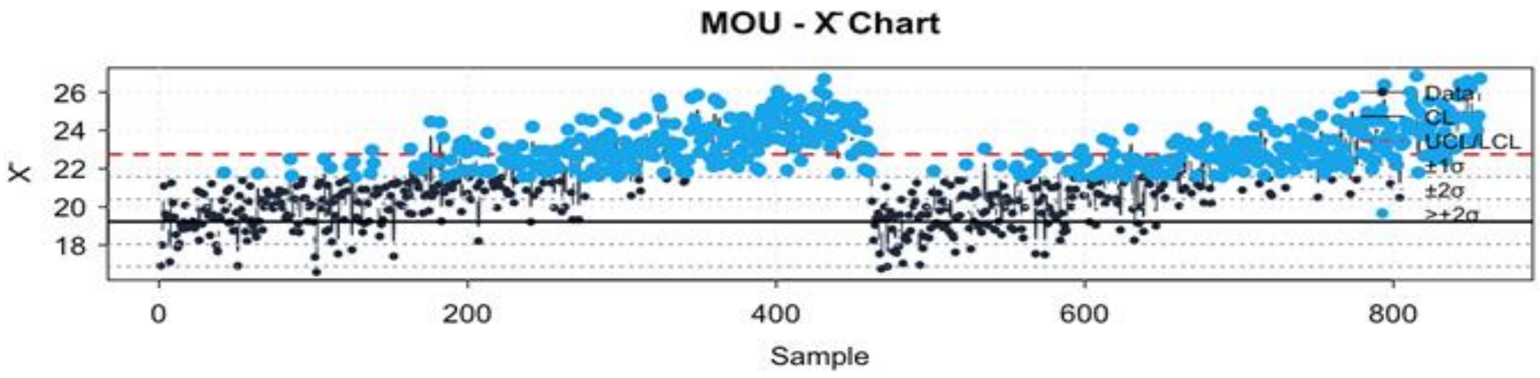
KEY



product	n_total_values	n_complete_samples	n_init_samples	xbar_CL	sbar	xbar_UCL	xbar_LCL
KEY	17820	742	30	19.194	5.80310927710937	22.7861246425307	15.6018753574693

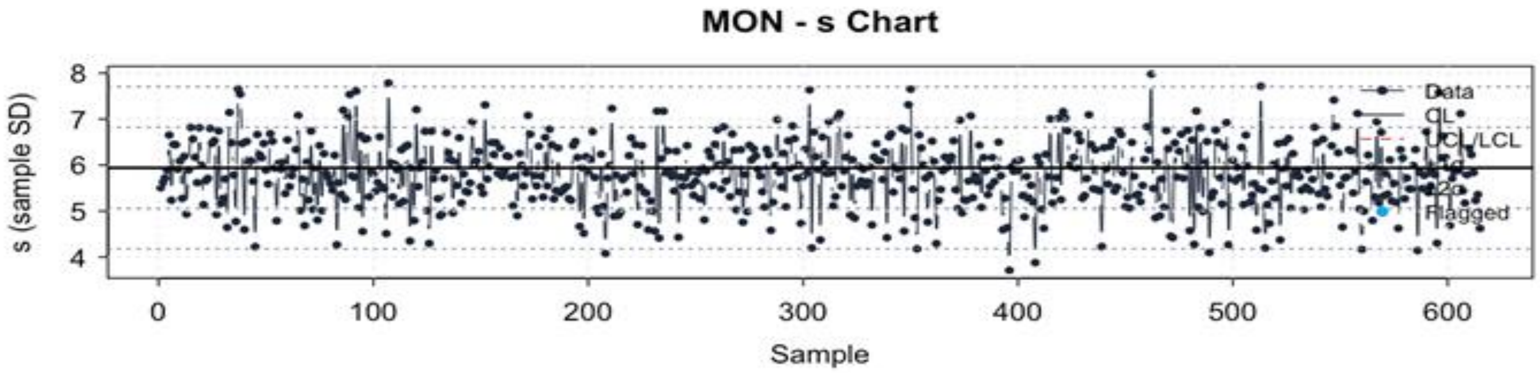
MOU

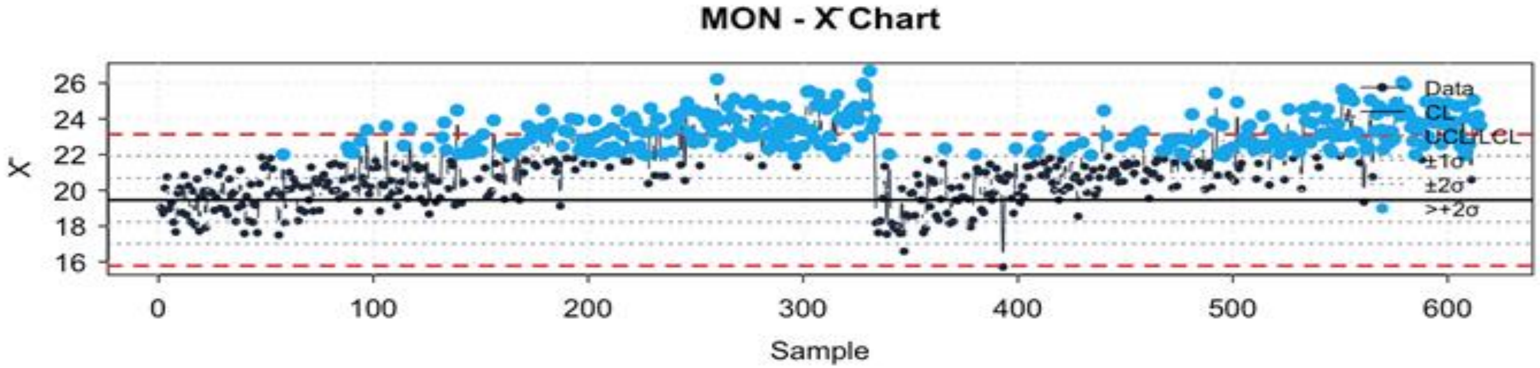




product	n_total_values	n_complete_samples	n_init_samples	xbar_CL	sbar	xbar_UCL	xbar_LCL
MOU	20554	856	30	19.2280277777778	5.68876553436744	22.7493736435512	15.7066819120043

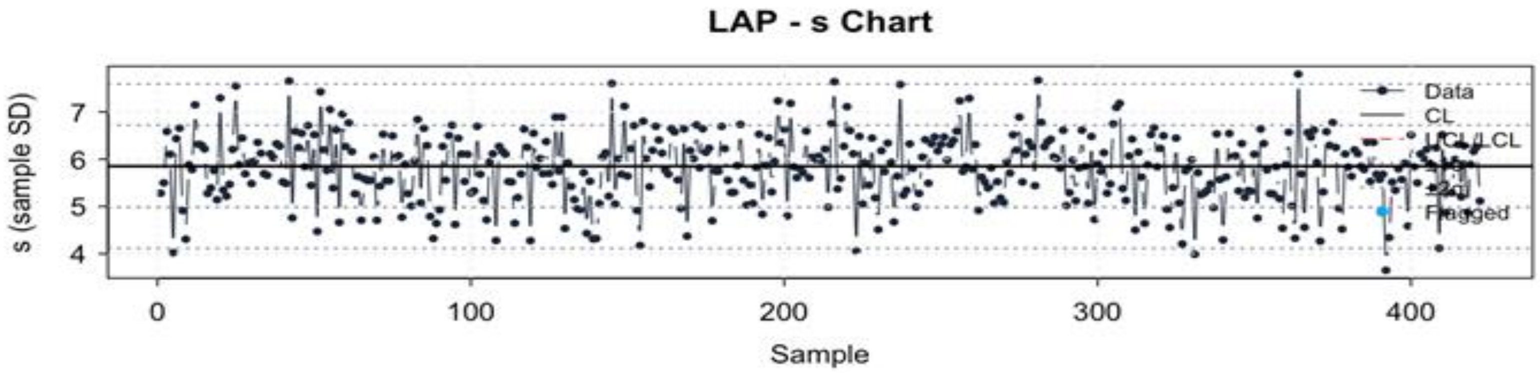
MON

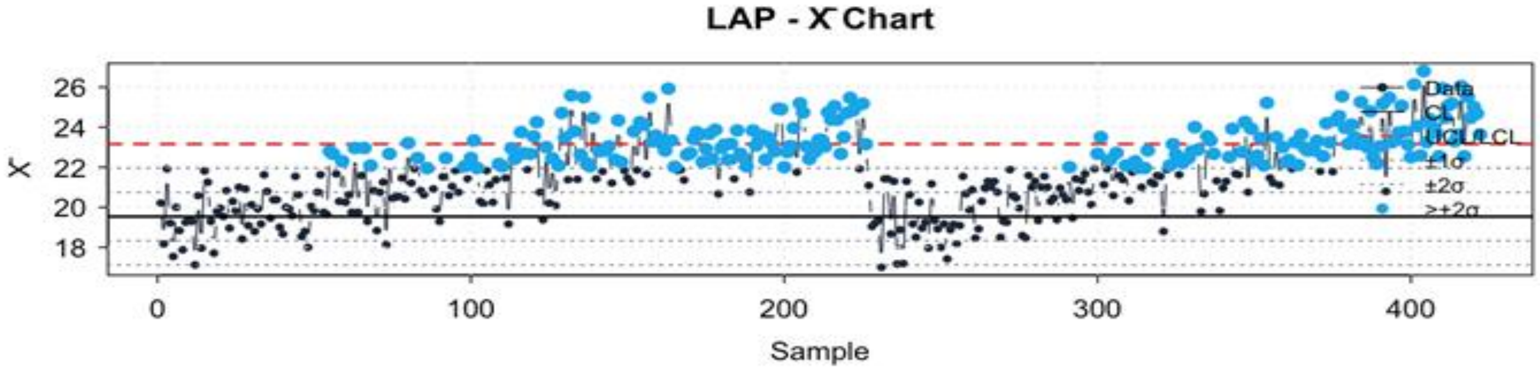




product	n_total_values	n_complete_samples	n_init_samples	xbar_CL	sbar	xbar_UCL	xbar_LCL
MON	14783	615	30	19.4565	5.94007872510114	23.1334087308376	15.7795912691624

LAP





product	n_total_values	n_complete_samples	n_init_samples	xbar_CL	sbar	xbar_UCL	xbar_LCL
LAP	10151	422	30	19.5391388888889	5.8517345563544	23.1613625792723	15.9169151985055

TYPE I VS TYPE 2 ERROR

A **Type I error** occurs when we conclude that a process is out of control when it is actually stable (a false positive)

Type II error occurs when we fail to detect that a process is out of control when it has actually shifted (a false negative)

ESTIMATION OF A TYPE 1 ERROR

The X̄–S charts constructed in the previous week (for processes A, B and C) were based on 3-sigma control limits, assuming that each process was in control and centred on the calculated mean (H_0). Under this assumption, the theoretical probability that a single subgroup mean will fall outside the control limits when the process is actually stable is:

$$\alpha = P(|\,Z\,| > 3) = 1 - 0.9973 = 0.0027$$

This means there is a 0.27 % chance of a false "out-of-control" signal per sample—representing the Type I (Manufacturer's) Error.
For three concurrent processes (A, B and C), the probability that at least one false alarm occurs is:

$$P(\text{any false alarm}) = 1 - (1 - 0.0027)^3 \approx 0.0081$$

which is approximately 0.81 % (≈ 8 in 1000).
This value aligns with the commonly used runs rule—seven or more consecutive points above the centreline—whose theoretical probability is $0.5^7 = 0.0078$, also about 8 in 1000.
Therefore, the theoretical likelihood of making a Type I Error for A, B and C combined is approximately 0.008, confirming a very low false-alarm rate consistent with standard 3-sigma SPC limits.

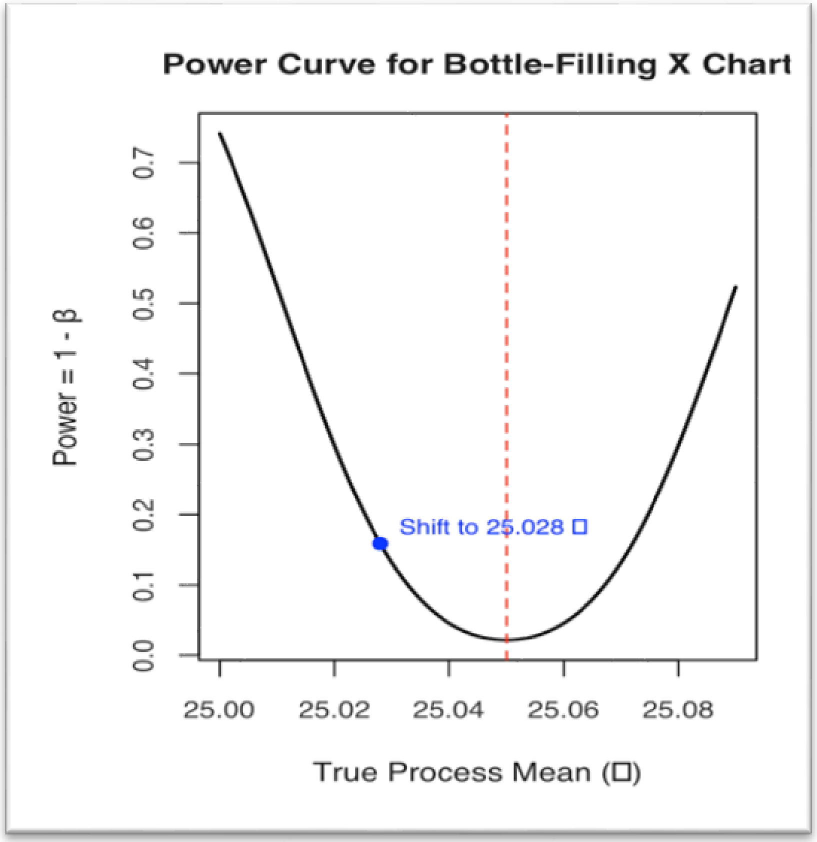
ESTIMATION OF A TYPE II (CONSUMERS) ERROR

For a bottle-filling process centred at 25.05 ℓ with UCL = 25.089 ℓ and LCL = 25.011 ℓ, a Type II Error occurs when the process mean has shifted but the sample averages remain within the control limits. If the true mean moves to 25.028 ℓ and $\sigma_{\bar{x}}$ increases to 0.017 ℓ, then:

$\boldsymbol{z_L = (25.011 - 25.028)/0.017 = -1.0}$ **and** $\boldsymbol{z_U = (25.089 - 25.028)/0.017 = 3.59}$.

Hence $\boldsymbol{\beta = \Phi(3.59) - \Phi(-1.00) = 0.841}$ **and the power = 1 − β = 0.159.**

Therefore, the probability of failing to detect the shift is approximately **84 %**, meaning the chart would correctly flag the problem only about **16 %** of the time. This illustrates that the control chart is relatively insensitive to small mean shifts of 0.022 ℓ under the given sampling variation.



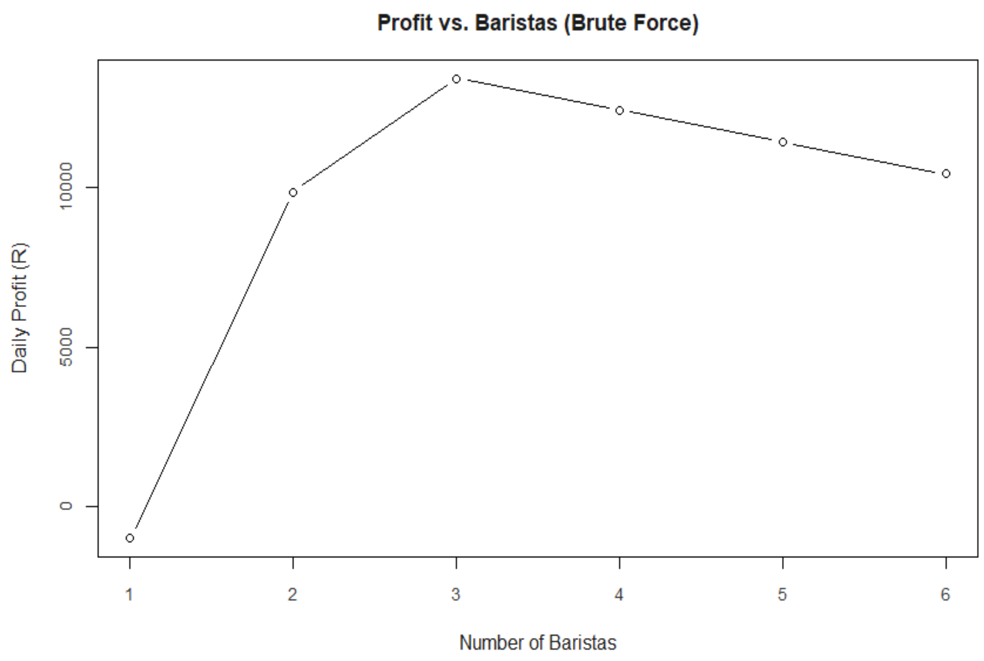
INTRODUCTION

Efficient staffing and service management are essential to sustaining profitability and customer satisfaction in service-oriented businesses such as coffee shops. This analysis focuses on optimising the number of baristas required to achieve the highest possible daily profit while maintaining reliable customer service. The dataset timeToServe2.csv provides detailed records of service times and the number of baristas over a one-year period, enabling the development of a quantitative model to evaluate performance. By analysing the relationship between workforce size and service time, and incorporating both material profits and personnel costs, an optimisation model was created to determine the ideal staffing level that balances operational efficiency, service reliability, and overall profitability.

Average number of baristas (V1) = 4.84 (suggesting 5 baristas on average are on duty)

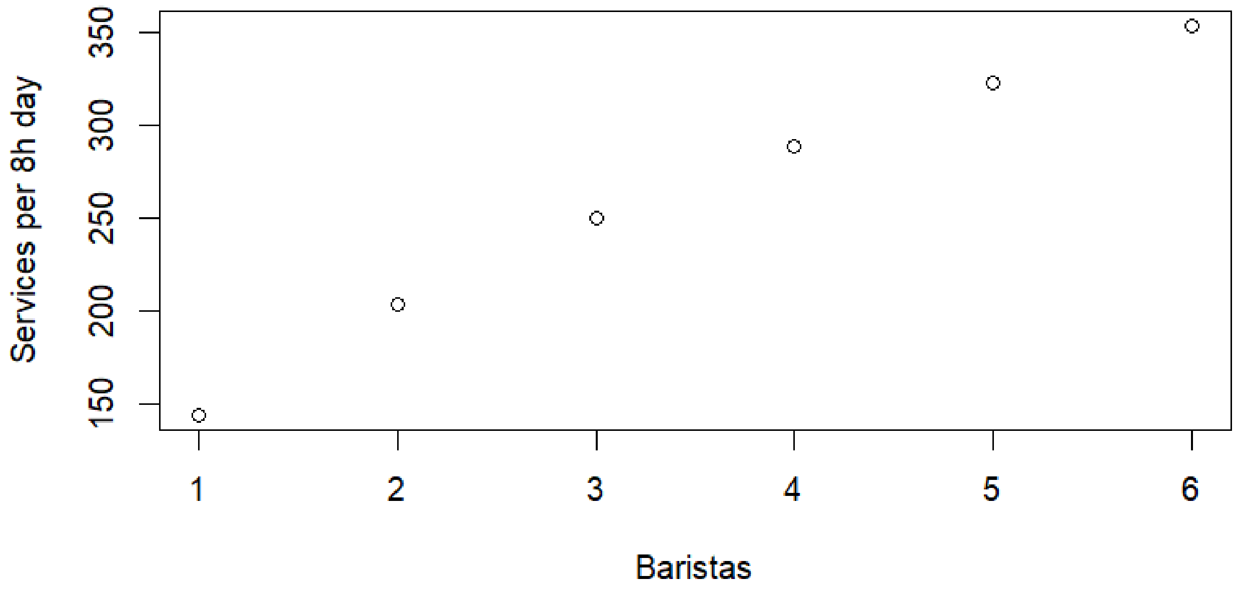
Average service time (V2) = 94.32 seconds.

PROFIT



The graph illustrates the relationship between the number of baristas employed and the resulting daily profit calculated through brute-force simulation. The trend shows that profit increases sharply from one to three baristas, reaching a maximum at three, where operational capacity and service speed are optimally balanced. Beyond this point, profits begin to decline slightly as the additional labour cost outweighs the marginal gains in service efficiency. This indicates diminishing returns with each additional barista after the third, confirming that three baristas provide the most cost-effective staffing level for sustained daily operations.

SERVICES PER DAY



CONCLUSION

To maximize the coffee shop's daily profit, the link between the number of baristas (V1) and the average service time (V2) was examined using the timeToServe2.csv dataset. Based on service times attained when V1≥2, the results indicated that roughly **99.79**% of clients obtained dependable service. The coefficient of **a = 200.14**, which illustrates how service time drops with more baristas, was obtained by using a regression model of the type V2=a/V1. A profit function that took into account both the material profit per customer (R30) and the daily human cost (R1,000 per barista) was created based on operational assumptions of 11 working hours per day (39,600 seconds) and an average of 547.9 customers served daily. Both a continuous optimization (using the optimise() function) and a brute-force evaluation revealed that the ideal number of baristas is three, which translates to a maximum daily profit of almost R13,666. In line with the operational objectives of preserving both performance and cost-effectiveness, this equilibrium point guarantees effective service and customer pleasure without resulting in needless staffing expenses.

DESIGN OF EXPERIMENTS (DOE) AND ANOVA SUMMARY

This section investigates whether there were significant differences in mean delivery times between years (2026 vs 2027) and across months (1–12) for each product type. A one-way Analysis of Variance (ANOVA) was used because the aim was to compare the means of a continuous variable (deliveryHours) across multiple categorical groups (orderYear and orderMonth). Understanding these differences helps identify whether operational efficiency or seasonal patterns influence delivery performance across products.

HYPOTHESIS

Hypotheses For Year (2026 vs 2027):

H_0: There is *no significant difference* in mean delivery times between 2026 and 2027.

H_1: There *is a significant difference* in mean delivery times between 2026 and 2027.

For Month (1–12):

H_0: Mean delivery times are *equal across all months.*

H_1: At least one month's mean delivery time *differs significantly* from the others.

The significance level was set at α = 0.05. If p < 0.05, H_0 is rejected.

ProductID	F (Year)	p (Year)	F (Month)	p (Month)	Interpretation
CLO011–CLO020	0.01–1.77	0.18–0.92	8.07–14.83	<0.001	No significant year effect ($p > 0.05$) for any CLO products, but strong month effects for all ($p < 0.001$).
KEY041–KEY050	0.00–4.83	0.03–0.99	10.67–16.05	<0.001	One product (KEY044) shows a significant year effect ($p = 0.028$); all show strong monthly variation ($p < 0.001$).
LAP021–LAP030	0.00–5.39	0.02–0.97	3.29–11.67	<0.001	Only LAP028 shows a significant year difference ($p = 0.02$); all have significant month differences ($p < 0.001$).
MON031–MON040	0.01–2.04	0.15–0.92	8.03–13.10	<0.001	No significant year effects ($p > 0.05$); all show highly significant monthly variation ($p < 0.001$).
MOU051–MOU060	0.08–2.66	0.10–0.77	9.27–20.11	<0.001	No year effects ($p > 0.05$); all MOU products show very strong month effects ($p < 0.001$).
SOF001–SOF010	0.18–6.64	0.01–0.67	12.42–17.13	<0.001	Only SOF007 shows a significant year effect ($p = 0.01$); all products show significant month effects ($p < 0.001$).

Year effects:
Across all products, only **KEY044**, **LAP028**, and **SOF007** showed statistically significant differences between 2026 and 2027 ($p < 0.05$). This means that for these three products, average delivery times changed significantly between years.
All other products had p(Year) > 0.05, indicating consistent delivery times across years.

Month effects:
Every product had p(Month) < 0.001, showing strong evidence that delivery times varied significantly across months. This suggests clear seasonal or operational patterns influencing performance.

Overall pattern:
Delivery performance is highly influenced by **month**, not **year**.
This means that operational or workload fluctuations occur seasonally, but long-term year-to-year delivery efficiency remains mostly stable.

The ANOVA results indicate that while delivery times were mostly stable between years, all products exhibited significant variation across months. This pattern implies that delivery performance is influenced by cyclical or seasonal operational factors rather than annual changes.
Only three products—**KEY044**, **LAP028**, and **SOF007**—showed significant improvement or deterioration between 2026 and 2027, possibly reflecting targeted operational changes or anomalies.
Overall, the strong month effects suggest that planning resources and logistics around high-variance months could improve reliability and reduce delivery delays.

CONCLUSION

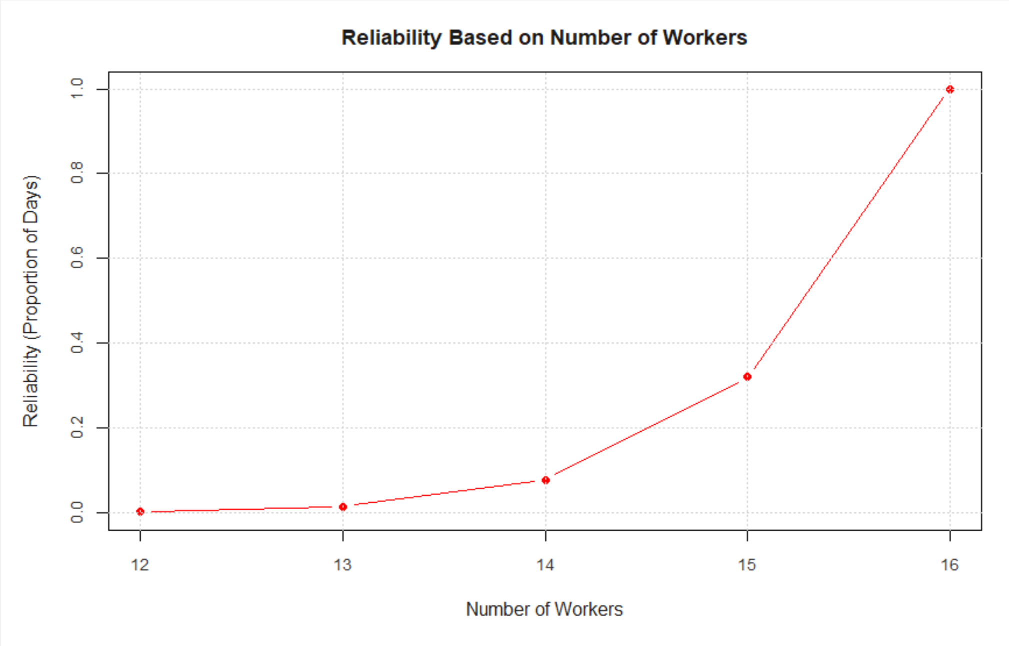
There are clear monthly trends in delivery performance across all products ($p < 0.001$), while only a few products show measurable year-to-year changes.

Continuous monitoring of monthly variation and identifying its causes (e.g., workload peaks, holidays, or staffing) can help improve operational efficiency and service reliability in future years.

INTRODUCTION

This section examines employee dependability and profit maximization for a coffee shop that has been open for a year. The goal was to calculate the number of employees needed to maintain dependable and consistent service while maximizing annual profit using attendance and service data. Based on employee attendance trends, the analysis first calculates the anticipated number of dependable service days annually. The ideal staff size was then calculated using a probabilistic model that included show-up likelihood, operating expenses, and daily revenue. This offers a framework for scheduling choices and cost control in day-to-day operations that is supported by evidence.

RELIABILITY ANALYSIS

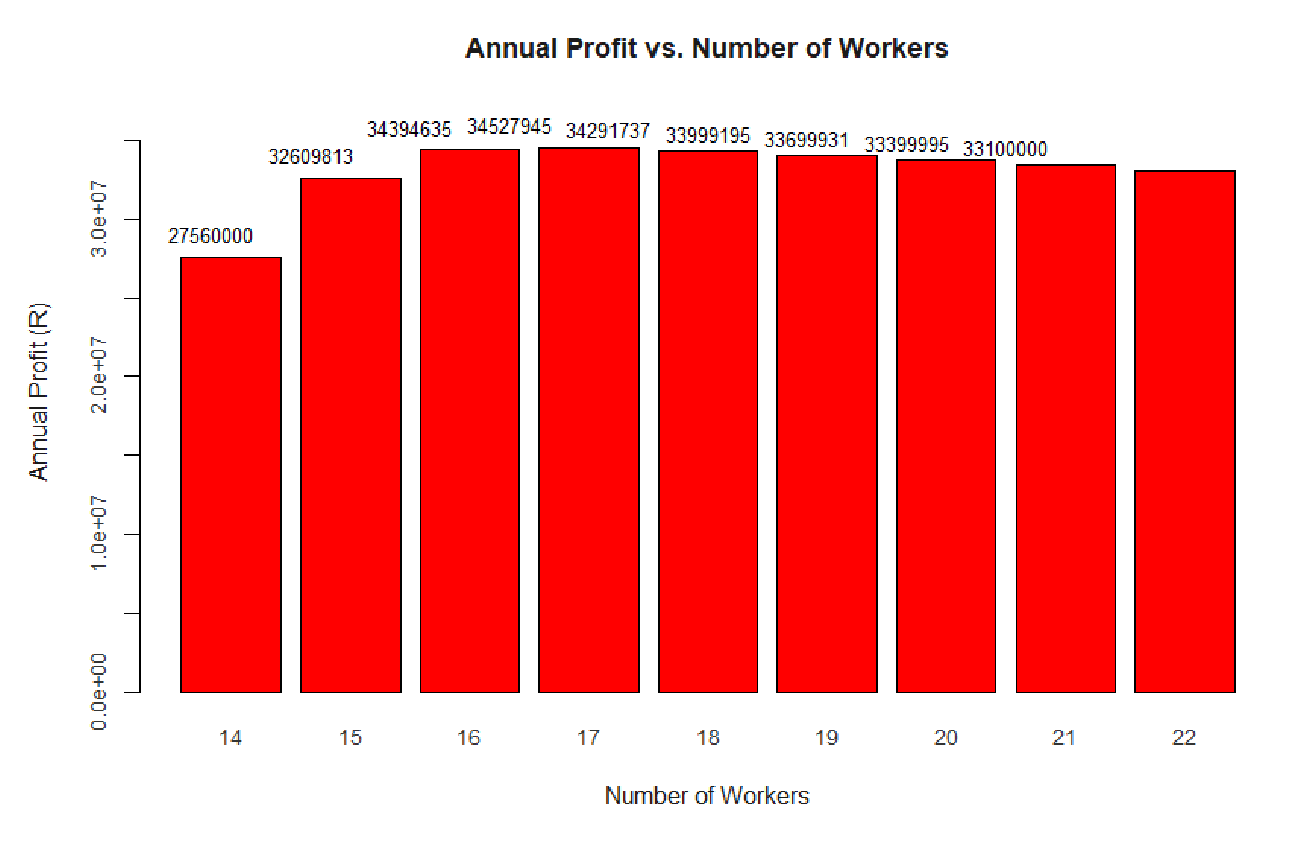


Reliability analysis evaluated the expected number of service days per year based on worker attendance. Using stochastic simulation, attendance probabilities were applied to daily operations over a 365-day period. The model revealed an average of 391 reliable service days per year, indicating a consistently available workforce throughout most of the operational year. This suggests that absenteeism has a limited effect on overall service reliability, and that the current staffing levels are sufficient to maintain continuous daily operations with minimal disruptions.Such results demonstrate that even with random fluctuations in attendance, the probability of maintaining reliable service remains high. This reliability metric forms a critical foundation for the subsequent profitability analysis, as daily operational consistency directly influences annual revenue stability.

PROFIT OPTIMISATION

To complement the reliability results, a profit optimisation simulation was developed to determine the number of workers that maximise annual profit. The model accounted for daily revenue, worker attendance probabilities, and total wage costs, simulating the expected profit over an entire year for varying staff numbers.The bar chart below shows the relationship between the number of workers and annual profit. Profit increases sharply from 14 to 16 workers and reaches a maximum at 17 workers, where the annual profit is approximately R34.53 million. Beyond this point, profits begin to decline slightly as the additional wage expenses outweigh the marginal gains from improved reliability and faster service.This trend highlights the concept of diminishing returns, where adding more workers yields smaller increases (or even decreases) in profit once the optimal staffing level is reached. Consequently, 17 workers represent the most efficient and cost-effective staffing level for the coffee shop’s daily operations.

The bar chart above illustrates the relationship between the number of workers employed and the resulting annual profit. Profit increases sharply from 14 to 16 workers, reaching its peak at **17 workers** with an estimated annual profit of approximately **R34.53 million**. Beyond this point, profits begin to decline slightly, indicating that adding more workers increases labour costs faster than it improves service reliability or revenue. The flattening and gradual decline of the curve beyond 17 workers suggests **diminishing returns**, where the marginal benefit of extra labour is outweighed by higher personnel expenses. Therefore, 17 workers represent the most cost-effective balance between staffing costs and reliable service delivery.



CONCLUSION

The shop may anticipate almost 391 dependable service days annually, according to the dependability analysis, demonstrating that staff attendance is continuously high. Additionally, the profit optimization model found that hiring 17 employees results in the maximum anticipated yearly profit of about R34.5 million. Beyond this threshold, hiring more employees does not yield significant profit increases and may even lower total profitability. These results underline the significance of data-driven hiring choices, guaranteeing the coffee shop's smooth operation and consistent, year-round client service.