Quality Assurance 344

ECSA Assignment

Final Report

K Folscher

26240076

Contents

Introduction

This project focuses on applying quality assurance tools to improve processes in an industrial engineering context. Using real-world data, several statistical techniques were used to understand patterns, monitor processes, and guide better decisions. The report starts with descriptive statistics to explore customer behaviour and product trends. It then uses statistical process control (SPC) methods—specifically X-bar and S control charts—to check if delivery times stay within expected limits. Next, the report calculates process capability indices (like Cp and Cpk) to see if delivery performance meets quality standards.

The project also investigates the chances of making Type I and Type II errors when monitoring processes and explains how reliable the control system is. After that, a Design of Experiments (DOE) and ANOVA are used to test if factors like year or month affect delivery times. Finally, the report looks at how staffing levels affect service reliability and business profit, helping to find the best number of employees for two shops. Each section builds on the last to show how data analysis can improve process quality and business performance.

Question 2

2) Descriptive Statistics

Means are tightly clustered in a narrow band (≈ \$80k–\$83k), indicating low between-city variation in average income. Miami and Chicago edge higher; New York and San Francisco are slightly lower. Income is homogeneous by city, so income alone probably doesn't explain large regional performance gaps.

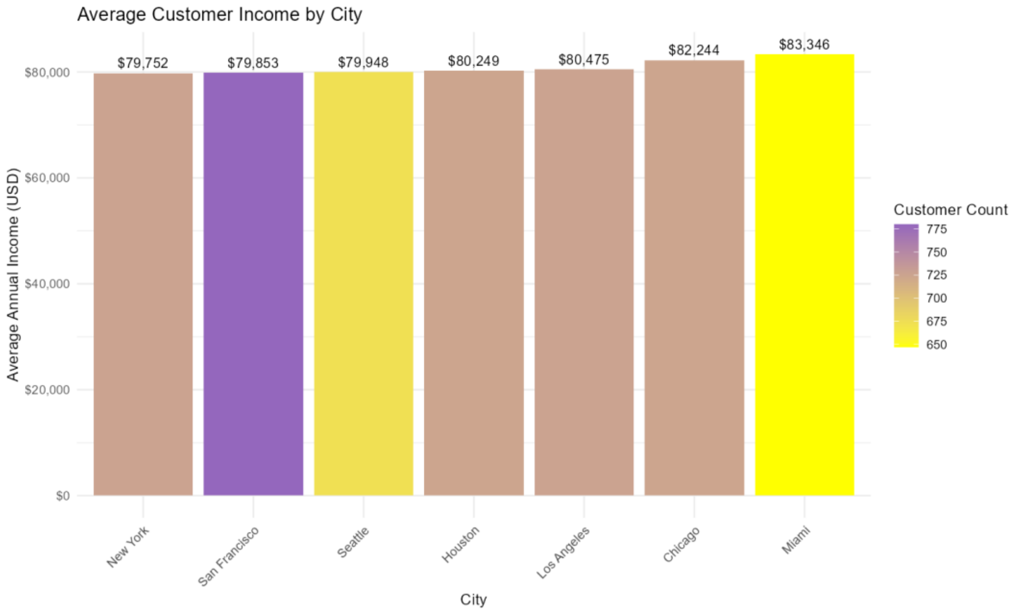

Figure 1: Mean customer income per city

Income shows substantial spread at every age (broad vertical dispersion), suggesting strong random/within-age variation and likely right-skewness (a few high-income points). Pearson r ≈ 0.16, indicating a weak positive linear association. Older customers earn slightly more on average, but age explains little variance in income (low r). Age is not a strong predictor of income in this sample.

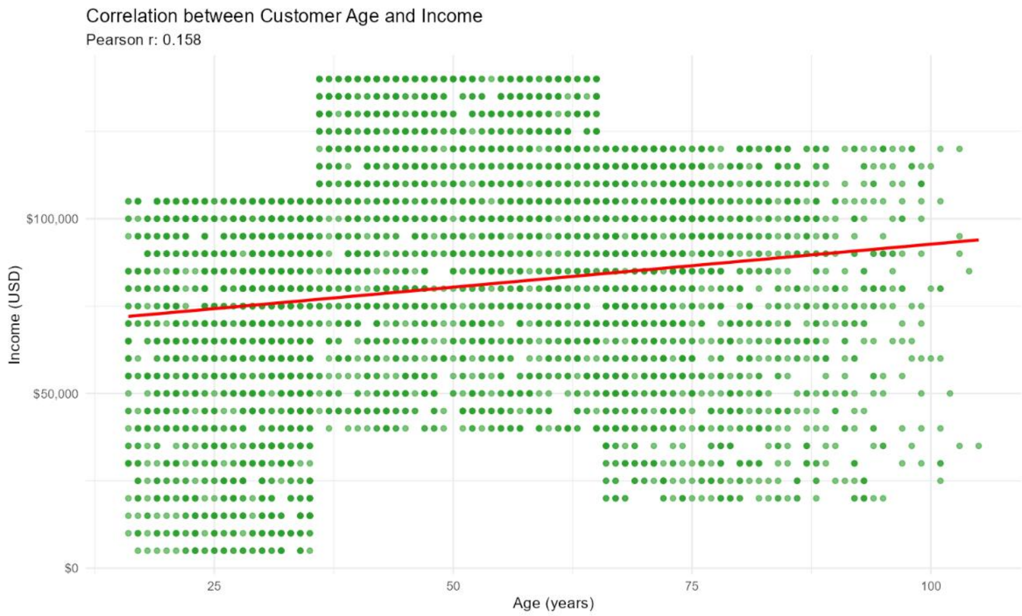

Figure 2: Scatter plot showing the relationship between customer age and income

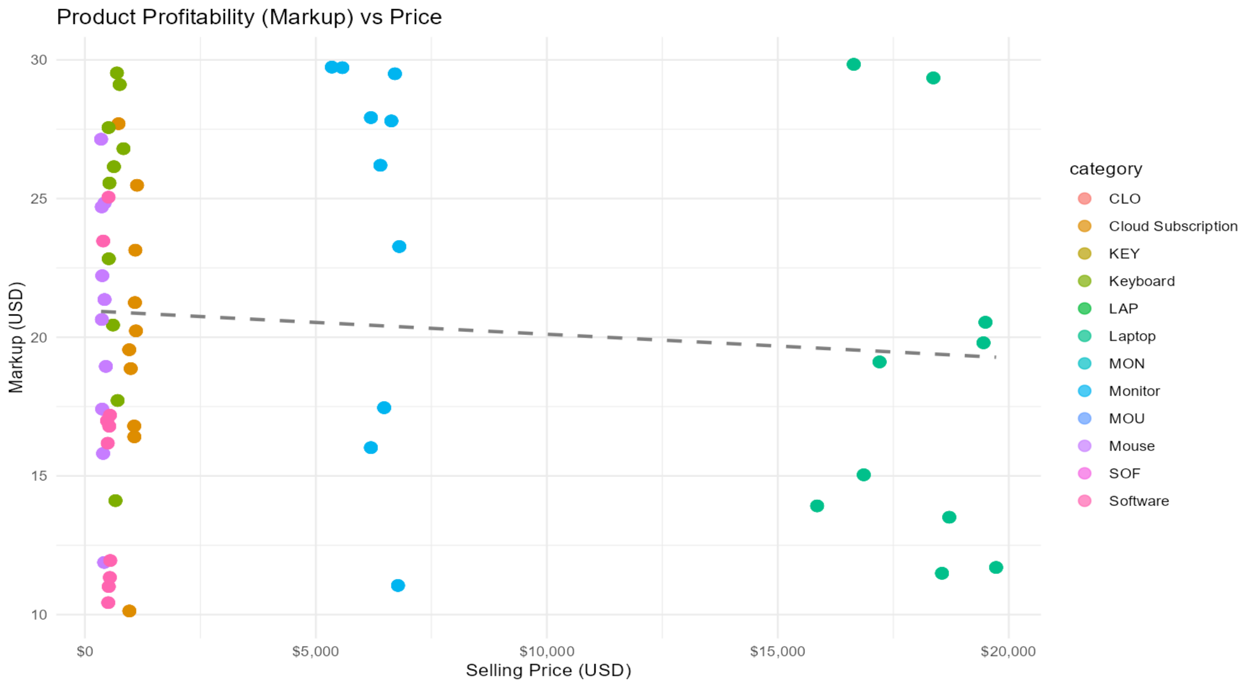

Figure 3: Product-level markup vs price by category

The fitted line slopes slightly negative → weak inverse relationship: higher ticket price does not guarantee higher dollar markup. Markups are concentrated in a narrow band (~$10–$30) while prices range from tens to many thousands—clear heteroskedasticity (spread in x much larger than in y). Category strategy matters more than price for gross profit per unit; focus on categories that achieve higher markups at sustainable prices.

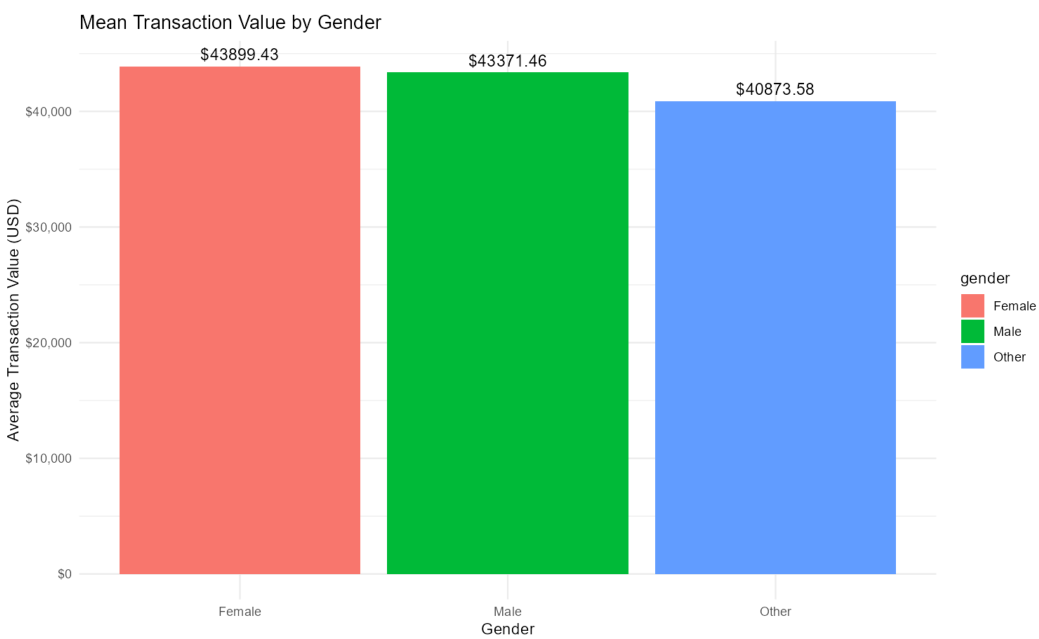

Figure 4: Average transaction value by gender

Group means are close, with Female ≈ Male and “Other” a little lower. Bars show only the mean; given typical right skew in basket values, group differences may not be statistically significant without a test. There is no material gap in purchasing power by gender; targeting should rely on other factors.

Percentages fall in a tight band (~22–24%), which is consistent with the threshold being global rather than city-specific. Miami is marginally higher and Seattle marginally lower. There is a very small between-city spread; city mix does not drive the proportion of affluent customers in a meaningful way. High-income concentration is uniform across cities; upsell plays should be national rather than city specific.

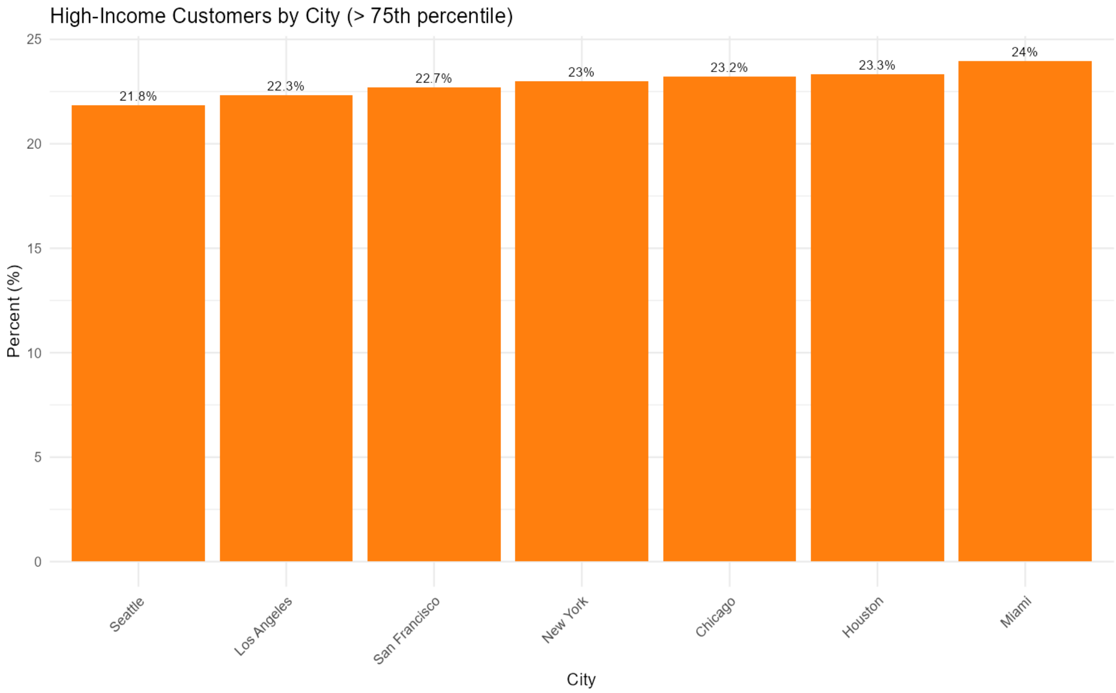

Figure 5: Share of customers whose income exceeds the overall 75th percentile

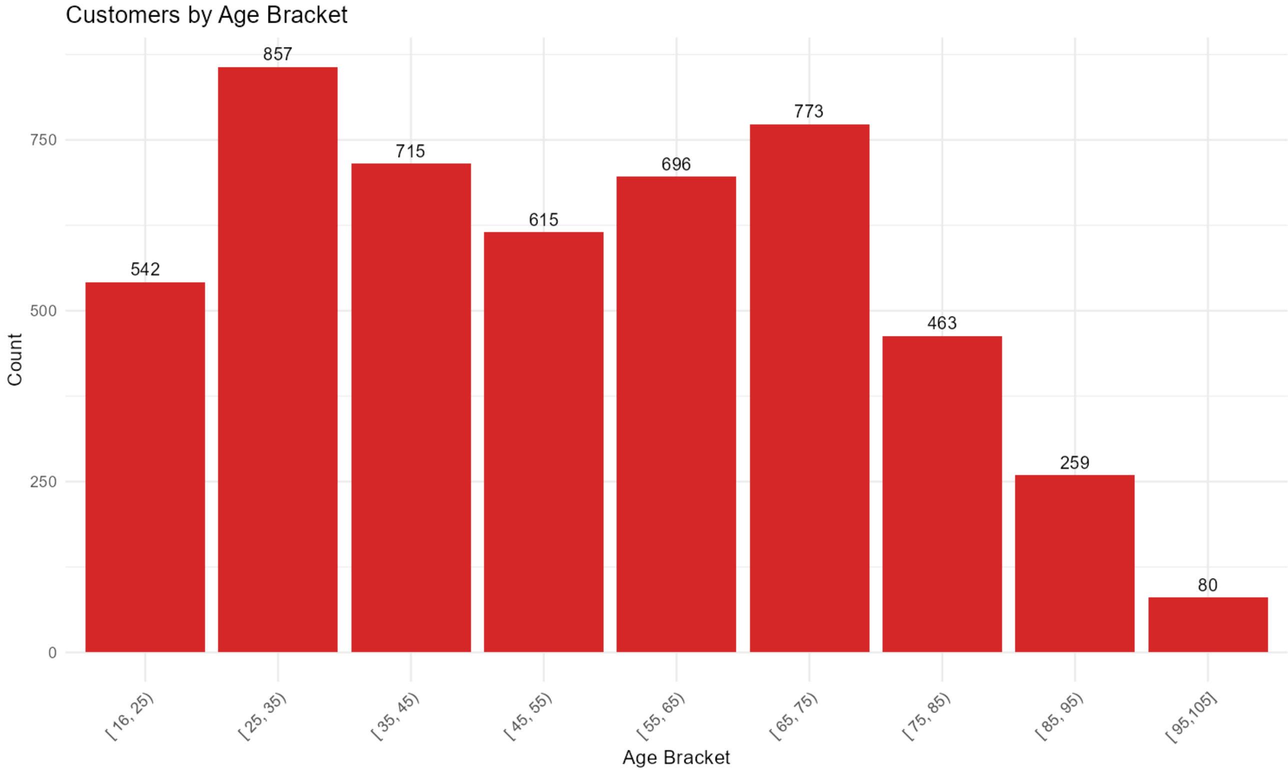

Figure 6: Frequency distribution of customer age

The count rises from 16–25 to 25–35, stays relatively high through 55–75, then declines after 75—i.e., a multi-modal / right-tailed age distribution with a long upper tail (older segments persist but shrink). The road coverage from teens to centenarians indicates wide population spread. The customer base skews toward working-age and early-retirement brackets. Campaigns should thus tailor to 25–75 with age-specific creativity.

Counts are stable within each year around a horizontal mean with minor random variation; large dips at year boundaries likely reflect partial months / data cut-off, not process changes. Month-to-month random fluctuations are small (no obvious seasonal cycle).

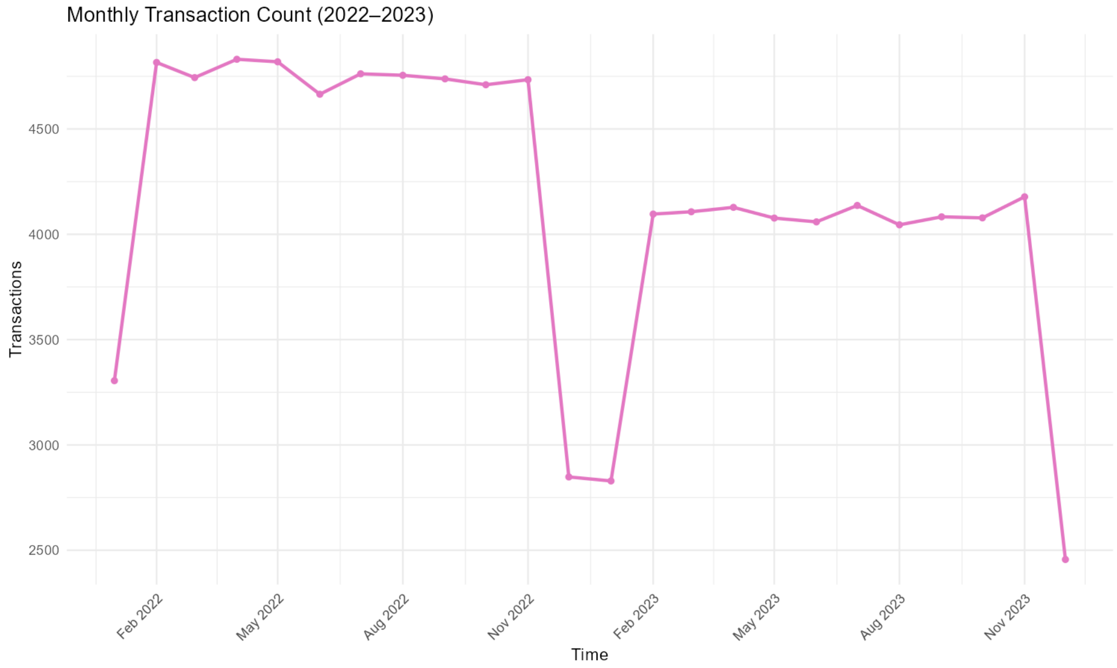

Figure 5: Time series of order counts

The demand volume is in-control and steady; resourcing (picking/fulfilment) can be planned at a near-constant level, adjusting only for month-end data artefacts.

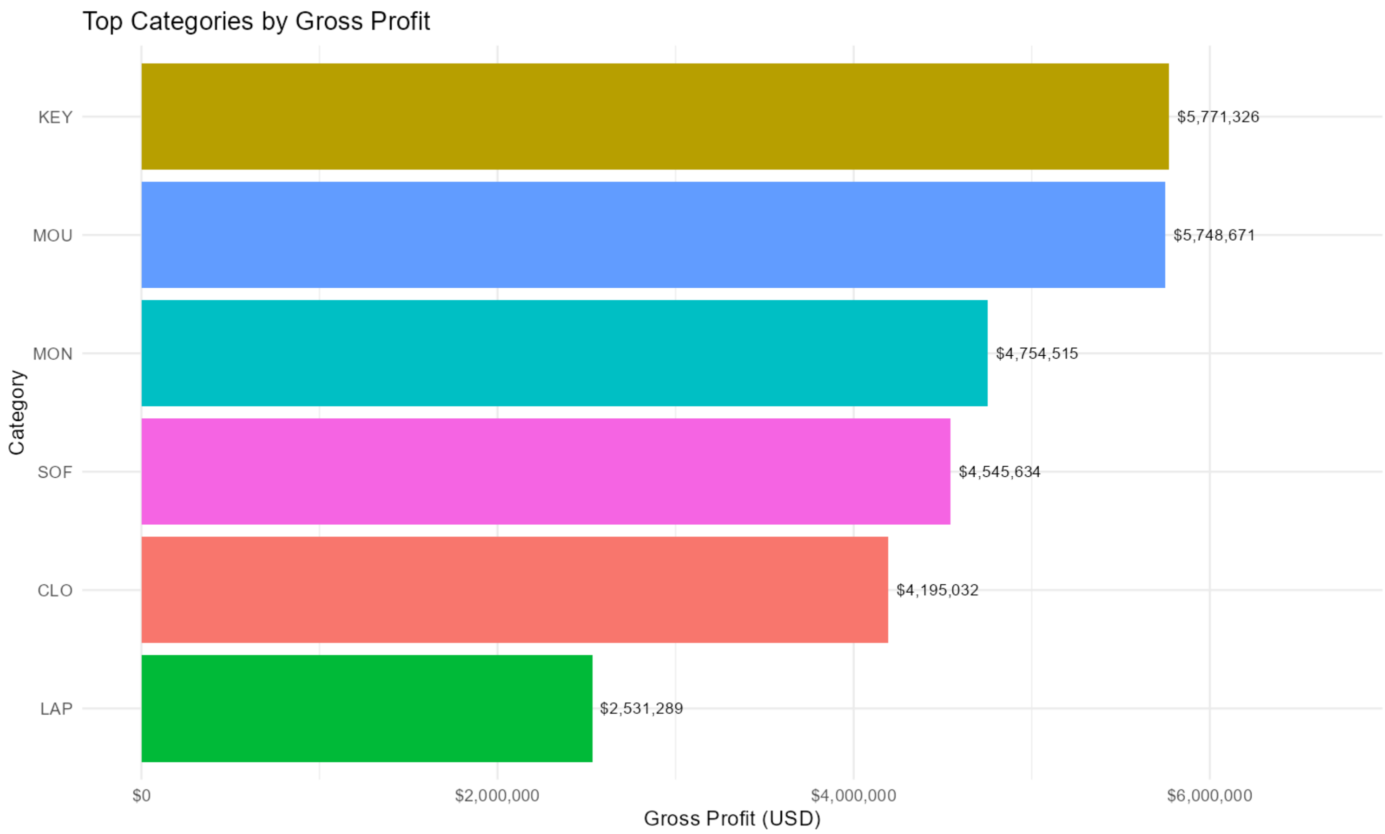

Figure 6: Total gross profit aggregated by category

KEY and MOU lead, followed by MON and SOF; LAP lags despite high unit prices—evidence that volume × markup outperforms high price alone. The bar lengths show meaningful between-category differences. Prioritise KEY/MOU/MON for promotions and inventory depth. Reassess LAP pricing/markup strategy or address conversion volume.

The curve is slightly increasing with a plateau mid-to-upper income—so higher income customers spend somewhat more per order, but the effect diminishes. There is very wide vertical dispersion at all income levels → weak predictive power. The transaction value is influenced by many factors besides income. The outcome is likely right-skewed; robust measures (median, trimmed mean) would complement the mean. Income has a non-linear, modest impact. Segmentation should combine income with behavioural features (recency/frequency/category preference).

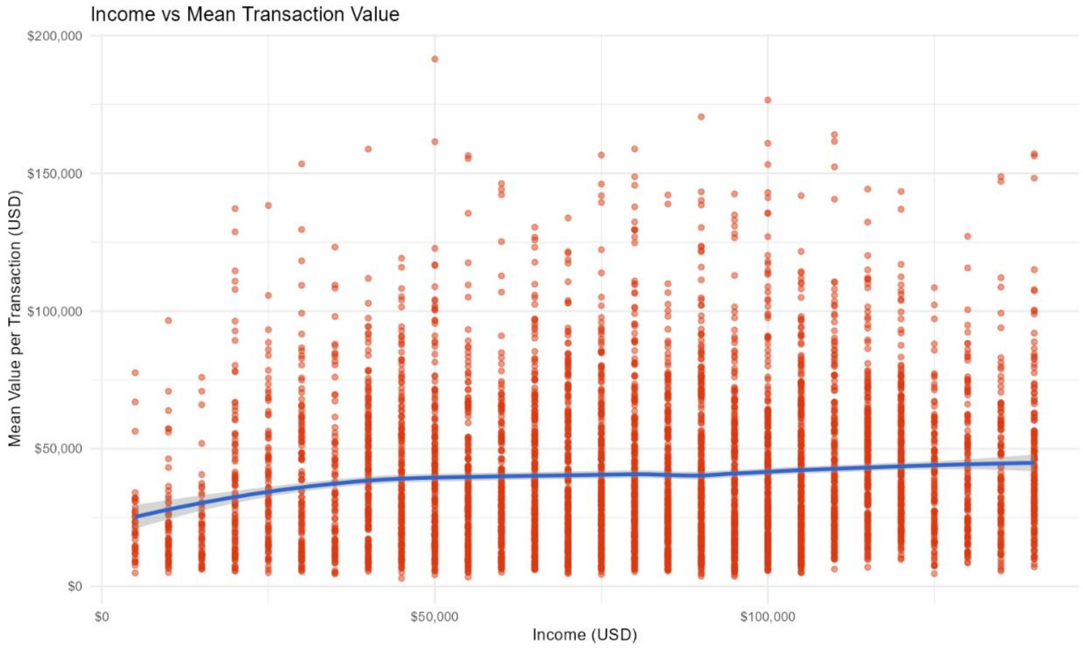

Figure 9: Per-customer income vs transaction value

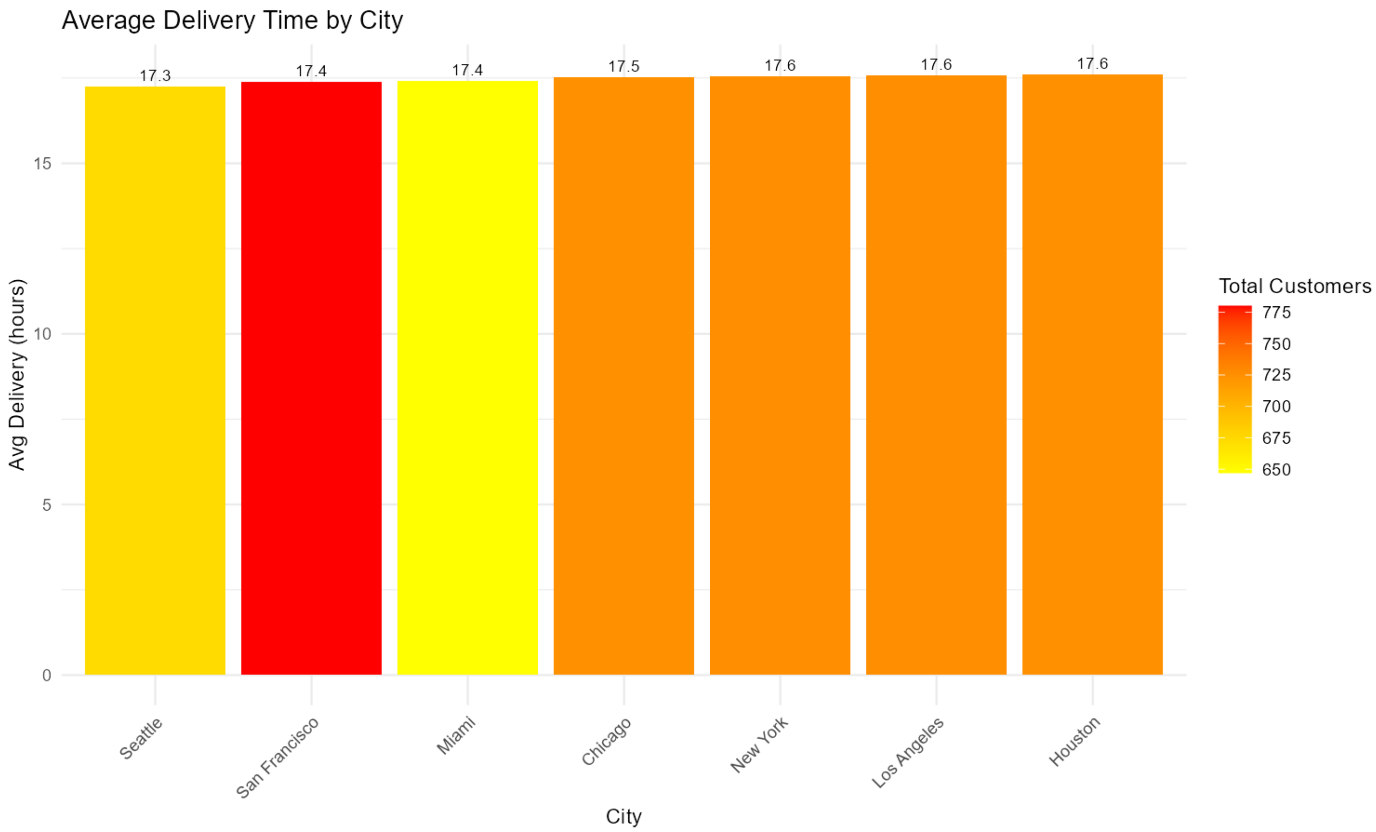

Figure 10: Mean delivery hours by city

Bars are tightly grouped around 17.3–17.6 hours. The narrow range suggests the delivery process is stable and centred across regions with random variation only; no city is outside practical control limits here. San Francisco and Houston are on the high side, and Seattle and Miami on the low side, but gaps are small (≈ 0.3 h). From a QA/SPC lens this is in control; improvements would require process-wide changes rather than city-specific fixes.

Summary of graphs

Across the updated datasets, descriptive statistics show homogeneous average incomes and high-income shares by city, a weak positive correlation between age and income, and a modest, non-linear relationship between income and average basket value with substantial random variation. Product-level analysis indicates markup does not scale with price; instead, category effects dominate, with KEY/MOU/MON delivering the highest total gross profit. Operationally, monthly order volumes appear steady without strong seasonality, and city-level delivery times are stable and in control, with only minor between-city variation.

Difference in outcomes

Revenue- and profit-based visuals (e.g., Profit vs. Price, Top Categories by Gross Profit, Mean Transaction Value by Gender) changed accordingly, while count/time metrics (e.g., Monthly Transactions, Average Delivery Time by City) remained stable. We also report 2023 total sales by product type from the updated price list. Any category misalignments present previously were removed by enforcing the 3-letter ProductID prefix mapping. (LAP = laptop; MON = monitor; CLO = cloud subscription; KEY = keyboard; SOF = software; MOU = mouse)

type	total_sales_value_2023
LAP	1163889479
MON	578385570
CLO	98715482
KEY	73499067
SOF	66468485
MOU	51219577

Table 1: Total sales values per product type

Question 3

3.1) X-bar and S Control Charts

Phase I:

Figures 17-22 display the first 30 samples of 24. The X-bar chart CL shows the average of the 30 samples mean delivery times and the S-chart CL shows the average of the 30 samples standard deviations. The X-bar and S-charts should have no strong trend or rule breaks. What was found:

Center (mean): For CLO, KEY, LAP, MON, and MOU, the X-bar points cluster around the center line with normal scatter inside ±2σ. SOF is centered near ~1 hour (a smaller-scale process).

Spread (variance): All S-charts stay between LCL and UCL; no points above +3σ. This shows stable within-subgroup variation in Phase I.

Signals: No meaningful SPC rule breaks. Phase I is in control and suitable for setting limits.

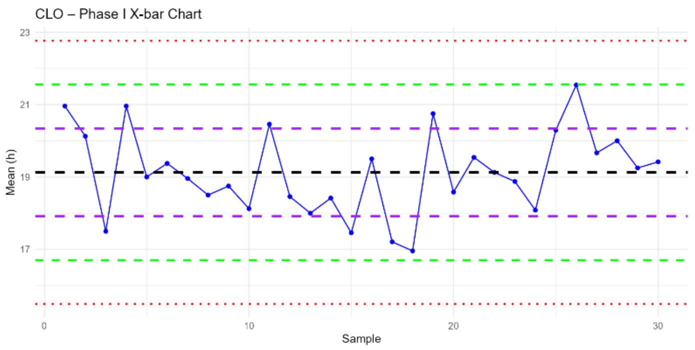

Figure 11: Cloud subscription X-bar chart

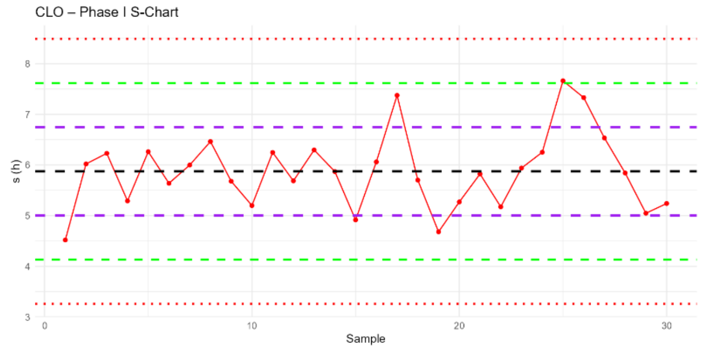

Figure 12: Cloud subscription S-chart

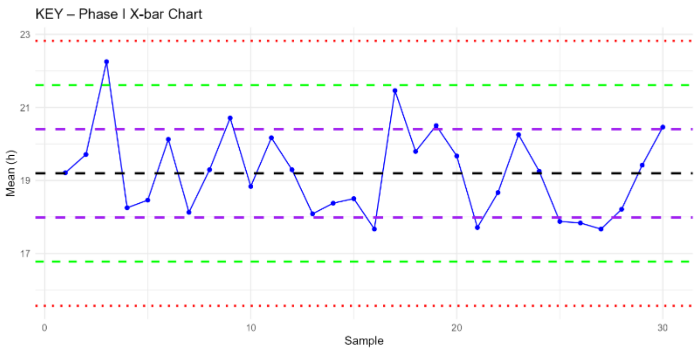

Figure 13: Keyboard X-bar chart

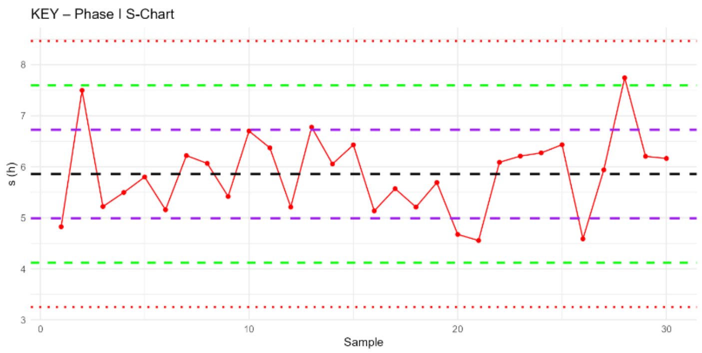

Figure 14: Keyboard S-chart

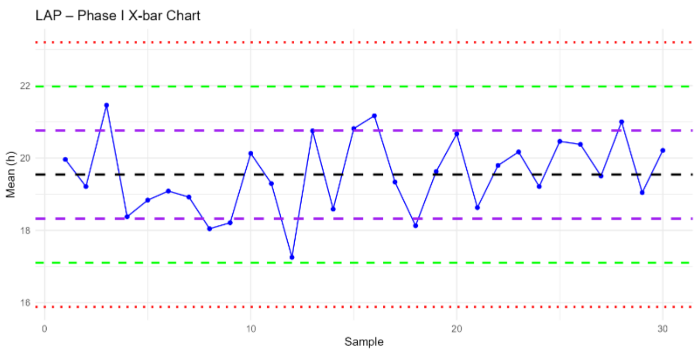

Figure 15: Laptop X-bar chart

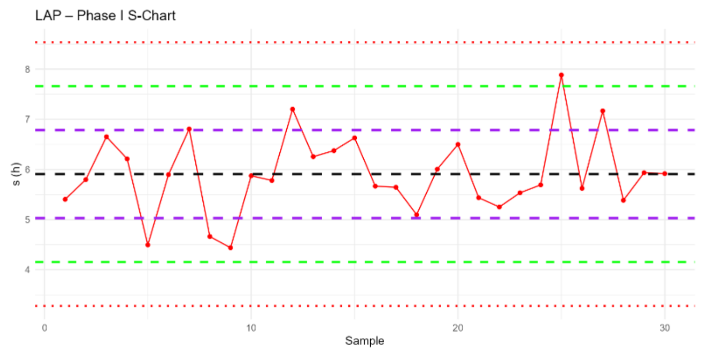

Figure 16: Laptop S-chart

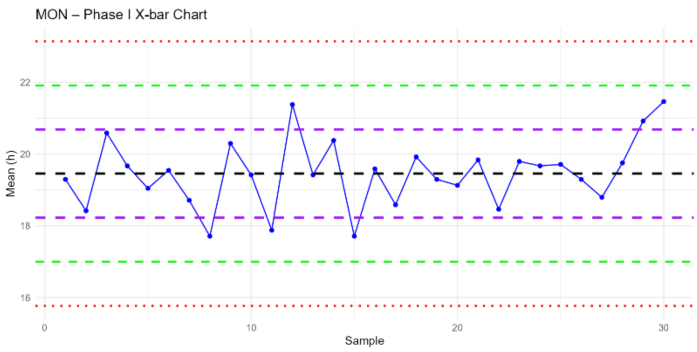

Figure 17: Monitor X-bar chart

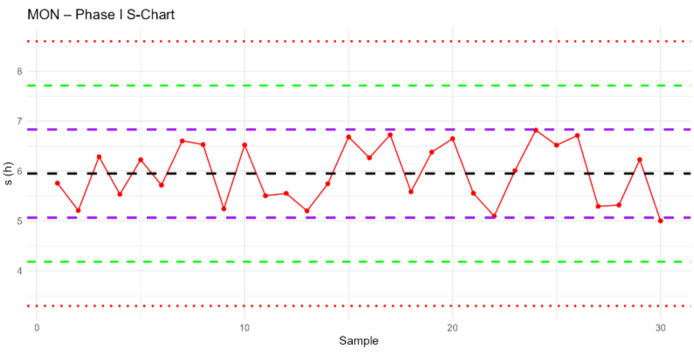

Figure 18: Monitor S-chart

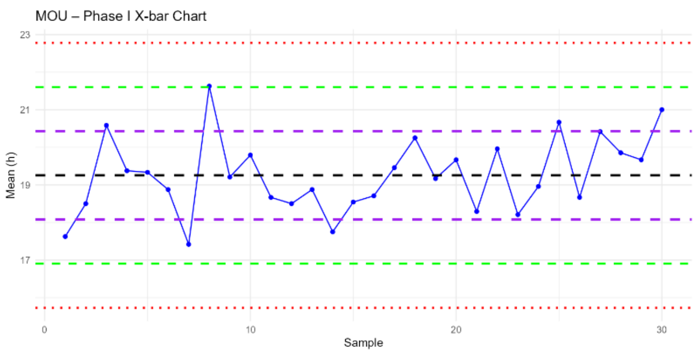

Figure 19: Mouse X-bar chart

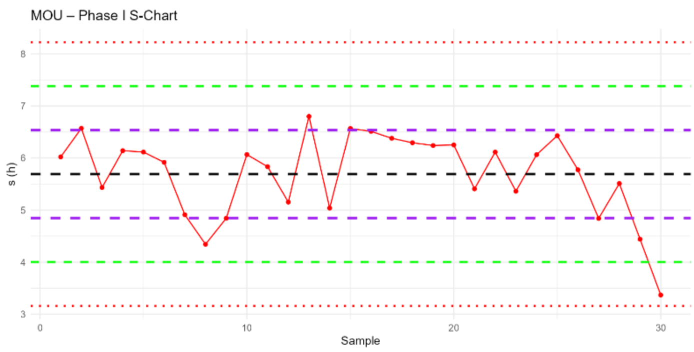

Figure 20: Mouse S-chart

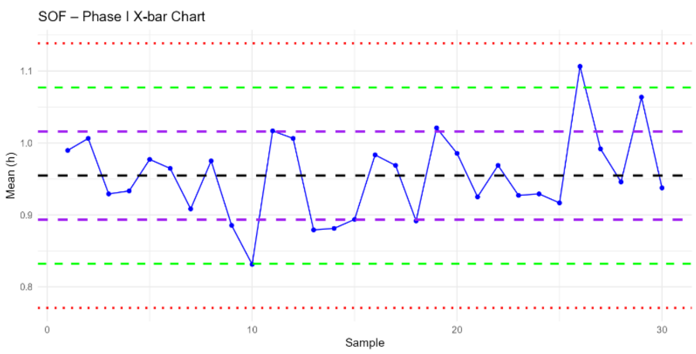

Figure 21: Software X-bar chart

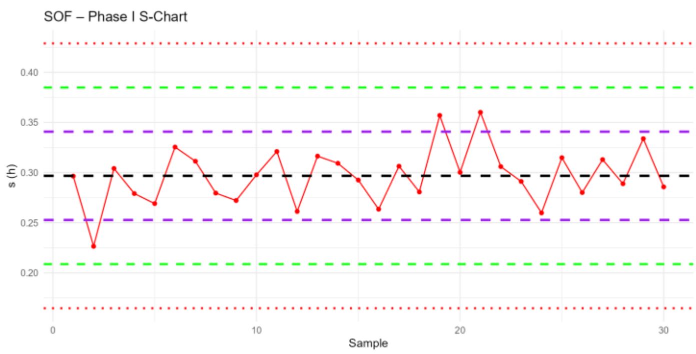

Figure 22: Software S-chart

3.2) Continuation of Sample Drawing

Phase II:

The following graphs (figures 23-28) display when the live process deviates from the basely, whether by shifting its average (X-bar chart) or getting noisier (S-chart). When there are many points hugging one side (above CL) then there is a shift in the mean. When the plots run near the CL the variation is stable and the control is good.

Rule A: when the spread blows past the upper limit on the S-chart, check for new suppliers, machine changes, different pickers/drivers/shifts, and batching/packing changes.

Rule B: if there is a long stretch of very stable variation, keep the settings, as this is desired.

Rule C: if there is a sustained upward shift on the X-bar chart, look for systemic causes such as routing changes or policy changes.

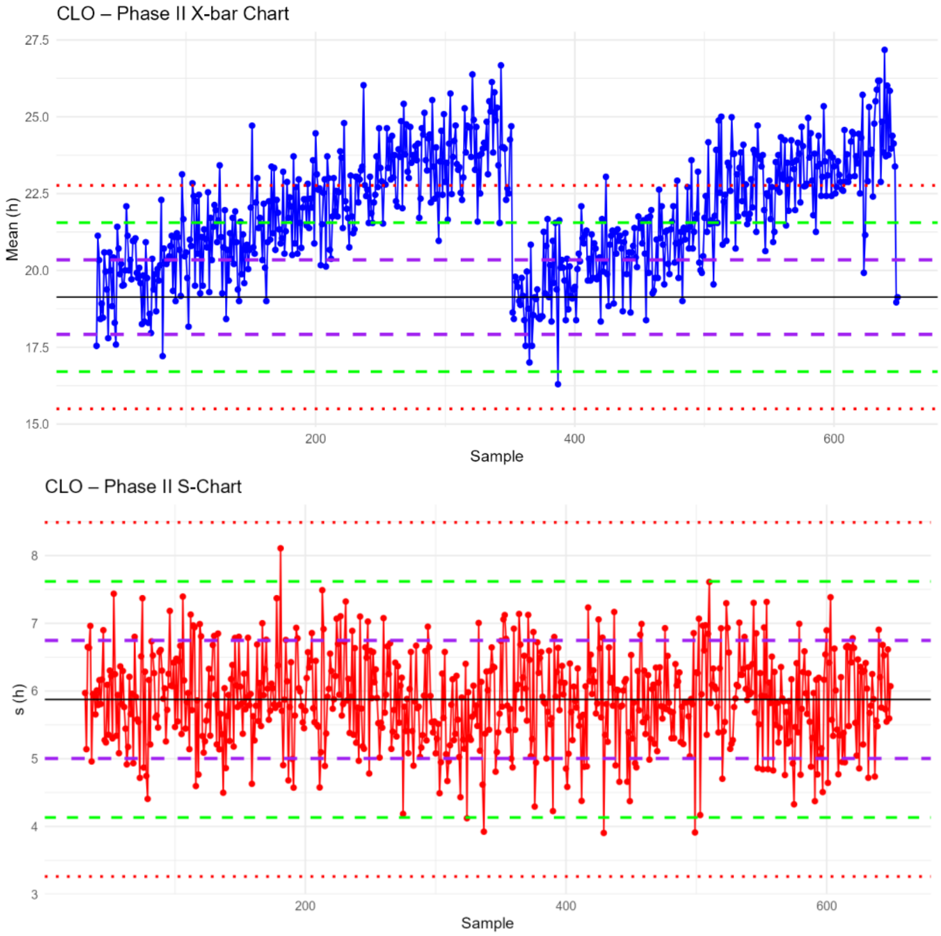

Figure 23: Cloud subscription X-bar and S-chart

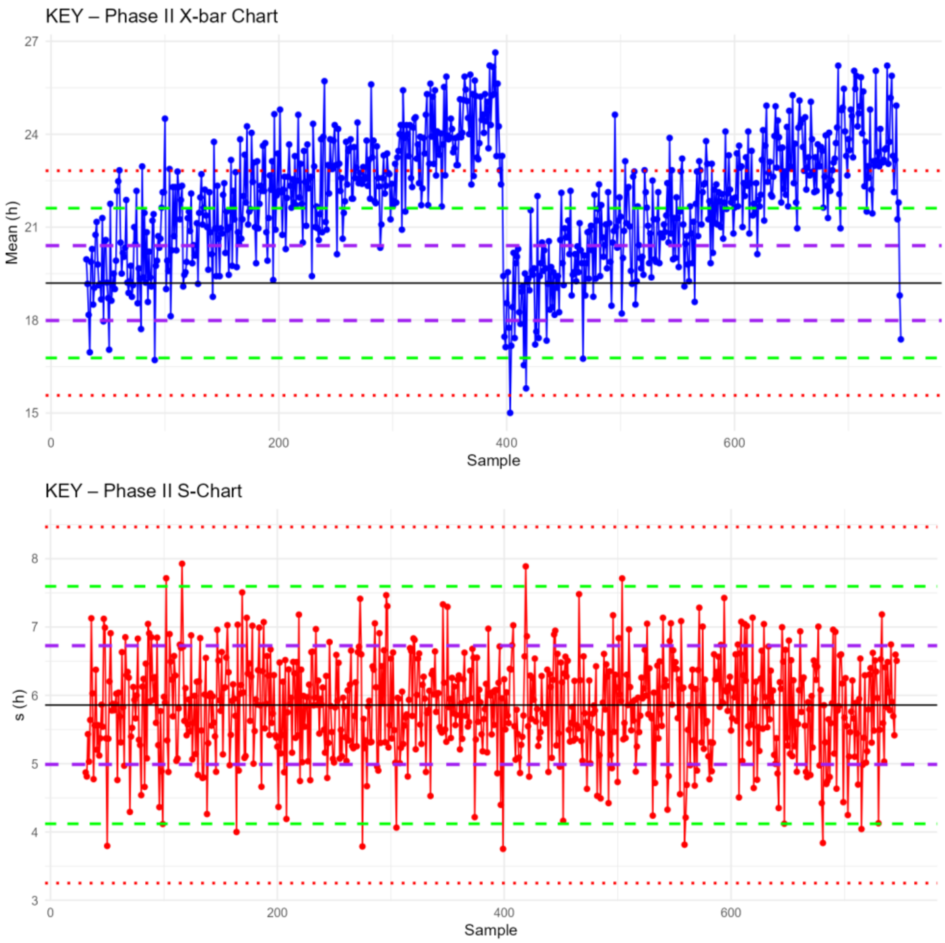

Figure 24: Keyboard X-bar and S-chart

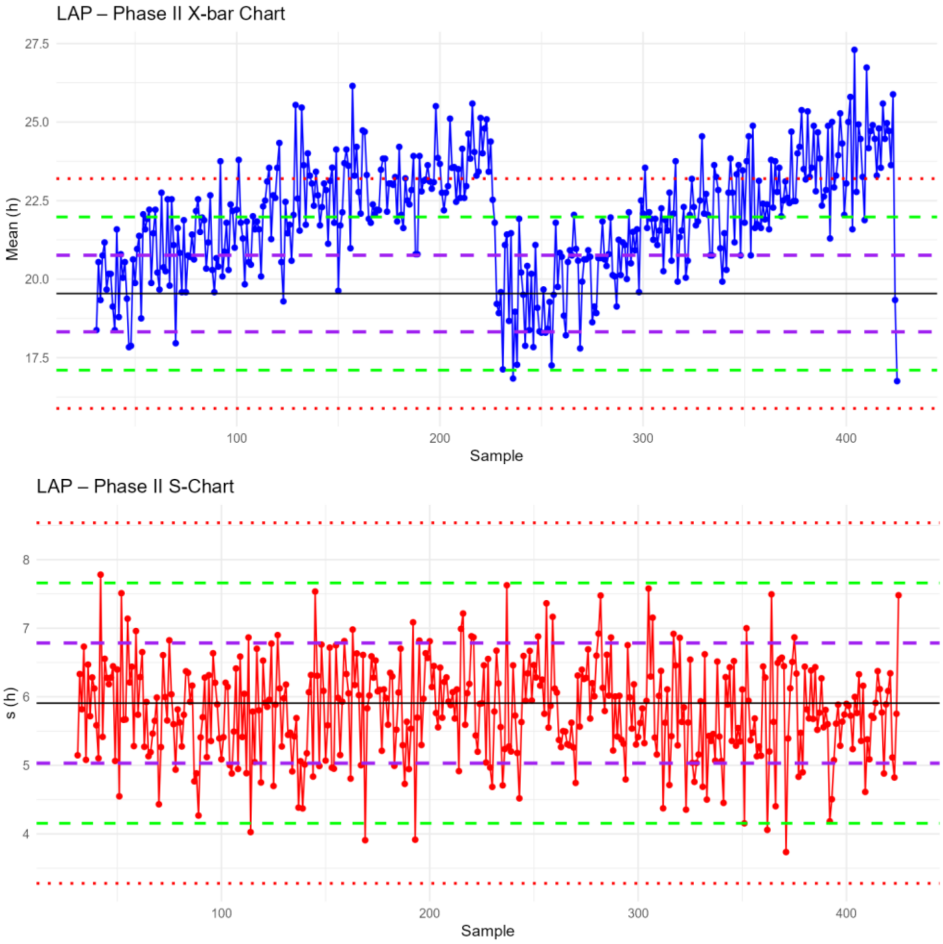

Figure 25: Laptop X-bar and S-chart

Figure 26: Monitor X-bar and S-chart

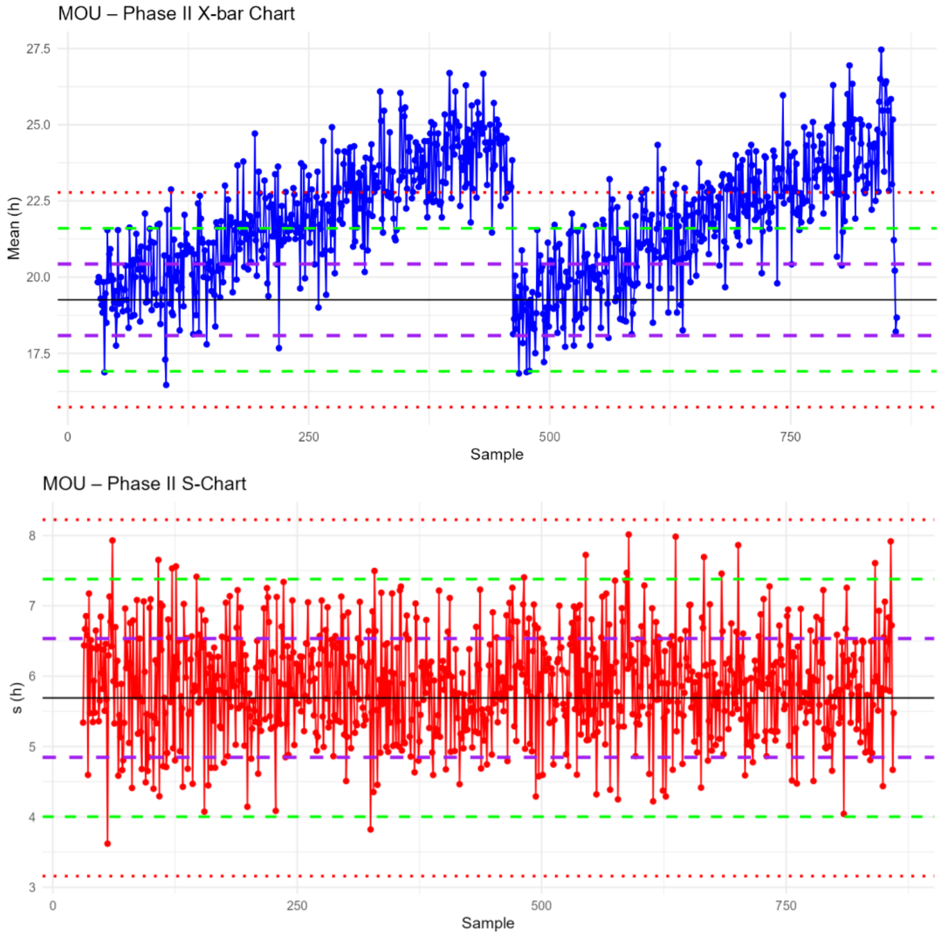

Figure 27: Mouse X-bar and S-chart

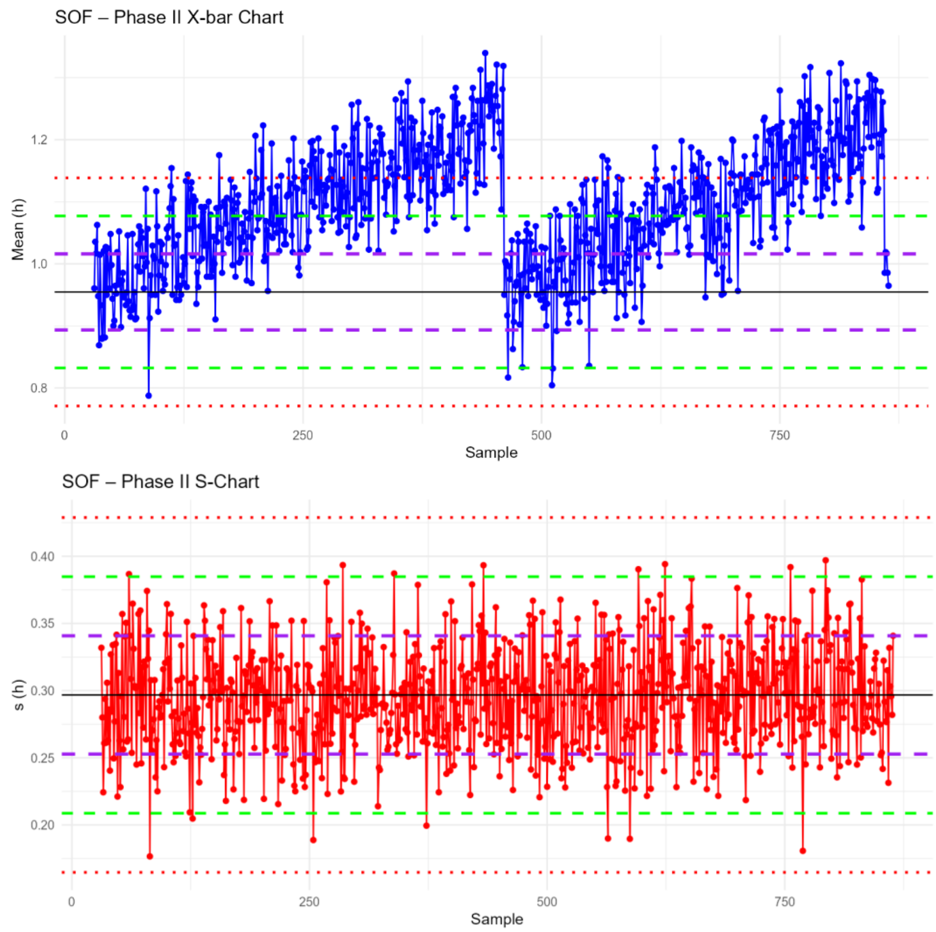

Figure 28: Software X-bar and S-chart

Across all product families, the means move (trends and step changes), while the variances stay stable. That is, the process centre shifts but spread does not.

- CLO (Cloud):

  The X-bar shows long upward trends, a clear dip around the middle, then a strong rise. There are several runs beyond +2σ (Rule C).

  The S-chart stays steady, meaning the variance in control.

- KEY (Keyboard):

  The X-bar has extended trends and a mid-series level shift; repeated Rule-C runs.

  The S-chart is stable and variance is unchanged.

- LAP (Laptop):

  The X-bar trends up, then there is a noticeable drop (~sample 200–250), then there is recovery and a rise. There are multiple >+2σ runs.

  The S-chart is within limits, meaning the spread is stable.

- MON (Monitor):

  The X-bar shows sustained mean increases and long runs.

  The S-chart is tight and inside limits; the variance is controlled.

- MOU (Mouse):

  The X-bar has several step changes and a later upward drift; Rule-C patterns are present.

  The S-chart is stable with no variance inflation.

- SOF (Software):

  The X-bar shows a drop then rise (clear level shifts) plus long trends. Despite a smaller scale, Rule-C conditions appear.

  The S-chart is low and steady; variance is stable.

3.3) Process Capability

mu: average delivery time over the first 1000 deliveries.

sigma: standard deviation (spread).

Cp: potential capability if perfectly centred.

Cpu/Cpl: capability to the upper/lower spec.

Cpk: the real capability (minimum of Cpu and Cpl).

VOC capable: capable if Cpk >= 1.33

product_family	mu	sigma	Cp	Cpu	Cpl	Cpk	n_used	Cpk_ge_1.00	Cpk_ge_1.33	VOC_capable
CLO	19.200	5.930	0.900	0.719	1.08	0.719	1000	FALSE	FALSE	Not capable (<1.33)
KEY	19.300	5.820	0.917	0.729	1.10	0.729	1000	FALSE	FALSE	Not capable (<1.33)
LAP	19.600	5.950	0.897	0.694	1.10	0.694	1000	FALSE	FALSE	Not capable (<1.33)
MON	19.400	6.000	0.888	0.699	1.08	0.699	1000	FALSE	FALSE	Not capable (<1.33)
MOU	19.300	5.830	0.915	0.726	1.10	0.726	1000	FALSE	FALSE	Not capable (<1.33)
SOF	0.956	0.294	18.100	35.200	1.08	1.080	1000	TRUE	FALSE	Not capable (<1.33)

Table 2: Process Capability

Interpretation of table 2:

If Cpk >= 1.33 then the VOC is comfortably met as per the project target. If Cpk is between 1.00 and 1.33 then there is a risk of misses when borderlines vary. If Cpk is below 1.00 the process is not capable and there is too much variation. When Cp is high and Cpk is low, the process could fit but is off-centre. The mean needs to be centred. When Cp and Cpk is both low the variation is too high, and sigma needs to be reduced. If Cpu < Cpl, then deliveries are too long or too slow.

3.4) Identify Samples that show Process Control Issues

This is the output of the rules for the process control:

=== CLOUD SUBSCRIPTION ===

Rule A (s > +3σ): total 0 | first/last: ...

Rule B (longest s within ±1σ): length 21 | samples: 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497

Rule C (≥4 X̄ over +2σ): runs 18 | 178–181; 200–203; 213–220; 222–225; 231–240; 245–251; 253–294; 296–341; 343–351; 518–523; 525–532; 534–539; 541–546; 548–552; 555–562; 564–585; 587–622; 625–647

=== KEYBOARD ===

Rule A (s > +3σ): total 0 | first/last: ...

Rule B (longest s within ±1σ): length 26 | samples: 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272

Rule C (≥4 X̄ over +2σ): runs 23 | 97–100; 169–172; 183–188; 198–202; 207–220; 225–228; 243–250; 252–255; 258–288; 291–297; 299–307; 312–396; 572–576; 594–598; 600–606; 608–612; 621–625; 627–654; 656–681; 683–692; 694–715; 717–720; 722–742

=== LAPTOP ===

Rule A (s > +3σ): total 0 | first/last: ...

Rule B (longest s within ±1σ): length 15 | samples: 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408

Rule C (≥4 X̄ over +2σ): runs 13 | 113–116; 118–121; 135–140; 152–155; 157–164; 167–178; 183–187; 190–226; 328–332; 363–391; 393–402; 404–408; 410–423

=== MONITOR ===

Rule A (s > +3σ): total 0 | first/last: ...

Rule B (longest s within ±1σ): length 14 | samples: 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75

Rule C (≥4 X̄ over +2σ): runs 23 | 181–184; 188–194; 201–204; 206–211; 213–217; 220–230; 237–242; 246–251; 254–270; 272–289; 291–294; 298–314; 316–333; 439–442; 480–483; 504–510; 518–521; 525–539; 547–560; 562–586; 588–596; 598–610; 612–616

=== MOUSE ===

Rule A (s > +3σ): total 0 | first/last: ...

Rule B (longest s within ±1σ): length 23 | samples: 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 824, 825, 826, 827, 828, 829, 830, 831, 832, 833, 834, 835

Rule C (≥4 X̄ over +2σ): runs 22 | 232–235; 247–250; 269–272; 278–295; 299–305; 312–326; 328–338; 342–439; 441–461; 661–664; 666–670; 677–680; 684–693; 698–706; 708–719; 722–727; 729–735; 737–750; 752–758; 760–797; 799–802; 805–856

=== SOFTWARE ===

Rule A (s > +3σ): total 0 | first/last: ...

Rule B (longest s within ±1σ): length 19 | samples: 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699

Rule C (≥4 X̄ over +2σ): runs 28 | 198–201; 205–208; 233–237; 240–244; 248–253; 257–266; 272–279; 281–285; 290–294; 300–304; 307–310; 322–328; 330–344; 349–360; 362–379; 381–387; 389–407; 409–421; 423–460; 644–651; 657–669; 698–705; 707–721; 723–728; 731–741; 746–757; 760–835; 837–859

Results interpretation:

Rule A:

No subgroup standard deviation exceeded the +3σ limit in the streaming phase (phase II). Thus, the variation is under control relative to the phase I baseline.

Rule B:

Long, good control runs on the S-charts show that the spread stayed very steady and the standard deviation is consistent. The focus should be on shifting the mean. Keyboard has the longest run with 26 samples inside ±1σ, Cloud Subscription = 21, Mouse = 23, Software = 19, Monitor = 14, and Laptop = 15.

Rule C:

There are a lot of runs above +2σ on the X-bar chart. The average delivery time increased for sustained periods relative to the Phase I baseline, which are systematic shifts. These are classic mean shifts with stable variation.

Question 4

4.1) Estimate the likelihood of making a Type I error for A, B, and C

Rule A and rule C Type I error are theoretical and constant per test. The probability of making a Type I error for A is 0.00135 (1 false alarm in 741subgroups) and for C is 2.68x10^-7 (1 false alarm in 3733054 4-point windows).

Rule B varies by family because it depends on the observed longest run and how many windows you scanned. Thus, the Type I (α) errors for B are:

Keyboard: longest run = 26; α = 0.033 (there is a probability of [α] that [longest run] runs are within one sigma of the centre line).
A 26-long streak inside ±1σ is rare even after scanning many windows; best control signal.

Mouse: longest run = 23; α = 0.117.
Still reasonably unlikely by chance; good control indication.

Cloud subscription: longest run = 21; α = 0.179.
Somewhat unlikely; moderate control indication.

Software: longest run = 19; α = 0.439.
Almost a 50/50 chance; weak control evidence.

Laptop: longest run = 15; α = 0.712.
Very plausible by chance; little to no special evidence of sustained control.

Monitor: longest run = 14; α = 0.937.
Essentially no evidence; the “best” run is exactly what you’d expect to happen randomly given many windows.

4.2) Estimate the likelihood of making a Type II error for a bottle filling process

A Type II (β) error is when the process has changed and the alternative hypothesis is true, but the test or chart fails to signal it, thus missing the problem. For the given scenario, the probability of making a Type II (consumer’s) error is β = 0.8412. This means there is an 84.12% chance of missing the shift ($\bar{X}$ stays between the LCL=25.011 and UCL=25.089) even though the true mean has moved to 25.028 L and the $\bar{X}$ SD grew to 0.017. The power is 0.1588 which indicates that there is only a 15.88% chance to detect this out-of-control condition on any given sample. The chart is thus very unlikely to catch the problem. To improve this, increase the subgroup size n, add supplementary rules, and reduce process variation at the source and re-estimate phase I limits before monitoring. The type II error is referred to as a false negative and it’s the cost of missing real problems. It is important to know since it tells you the risk of quietly drifting out of control.

Question 5

Assumptions:

- 10 hour working days and 24 working days per month
- Material profit = R30 per customer and barista cost = R1000 per day
- Reliability metric: % of customers served within 120 seconds
- Caps baristas at 6
- Evaluate using individual service times

Method:

- Capacity model: For each shop the mean was summarised and 95th-percentile service time per barista count, then fitted a simple curve
  Mean service$(B) = a + b\left(\frac{1}{B}\right)$ to estimate orders/day = $\frac{10\times3600}{\text{Mean service}(B)}$.
- Demand: Estimated working-day demand from the file size (rows ÷ 288).
- Profit: Profit/day $= \min$ (Demand,Capacity) $\times 30 \; - \; (B \times 1000)$,
  then rolled up to monthly and yearly.

Results:

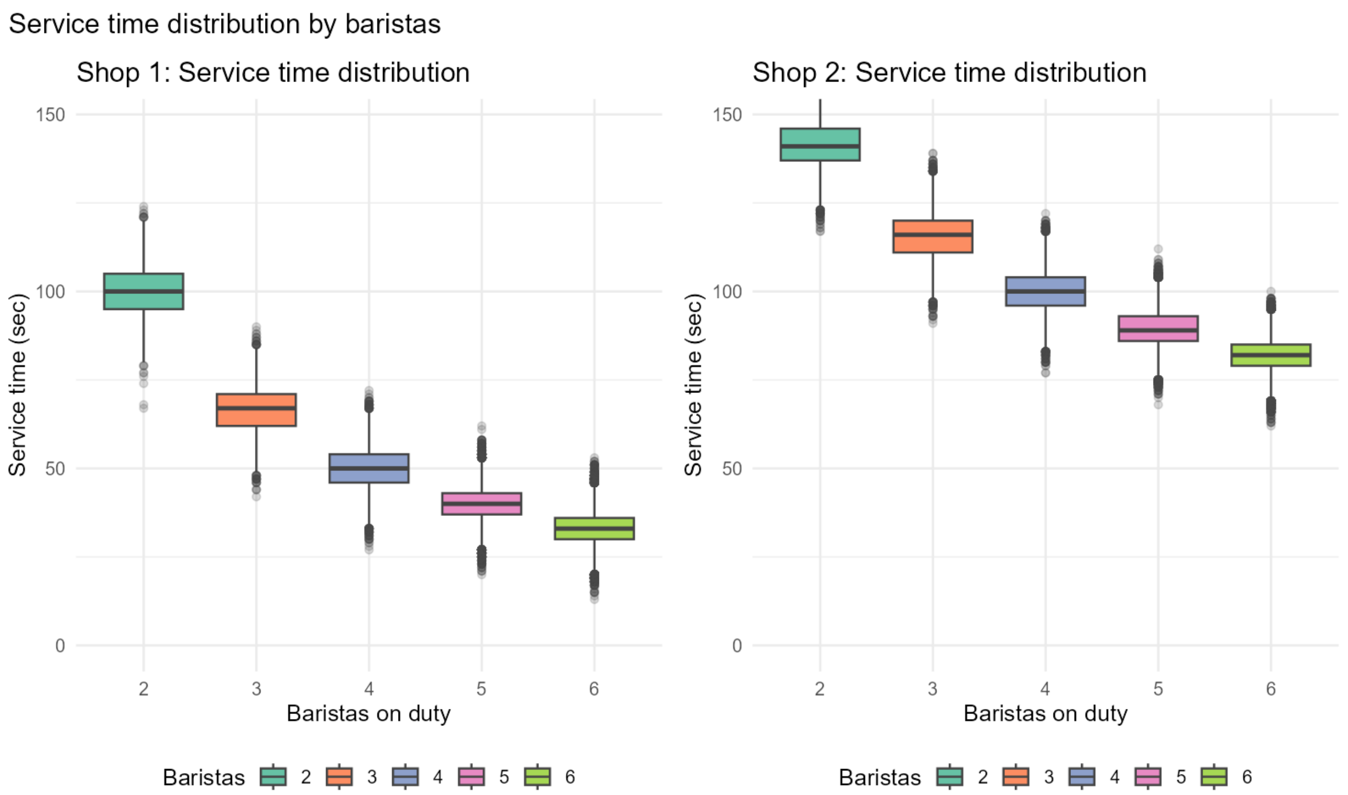

Figure 29: Boxplots of service time distribution by baristas

The boxplots show how the median and spread of service times change with staffing. As baristas increase, both shops show lower medians and tighter interquartile ranges, i.e., faster and more consistent service. Shop 1's distribution is generally lower (faster) than Shop 2's at the same staffing levels. Remaining outliers shrink as staffing rises, indicating fewer exceptionally slow orders.

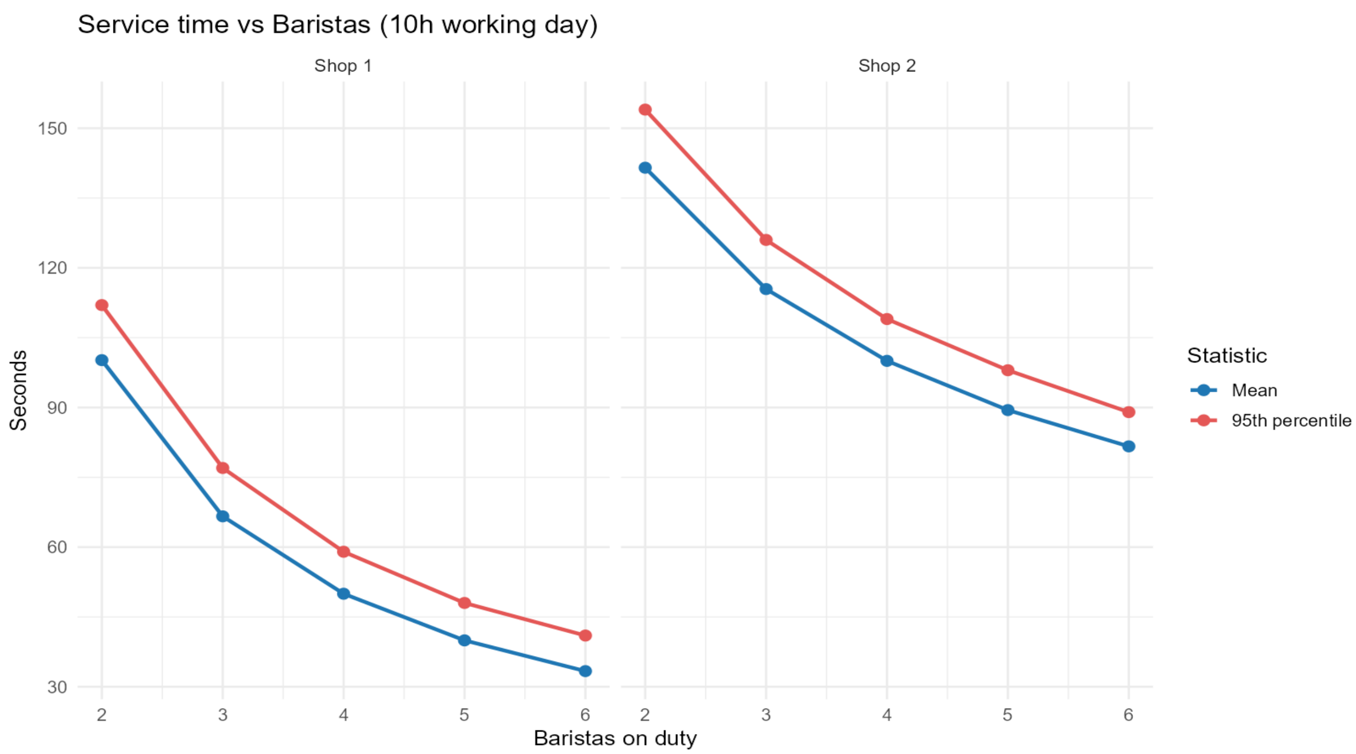

Figure 30: Service time per number of baristas per shop

Figure 30 shows that average and 95th-percentile service times decrease monotonically with staffing in both shops. The 95th-percentile line demonstrates reliability: with more baristas, nearly all customers are served within 120 seconds. Shop 1 benefits strongly up to 4 baristas, while shop 2 improves steadily but with smaller marginal gains, suggesting different process constraints.

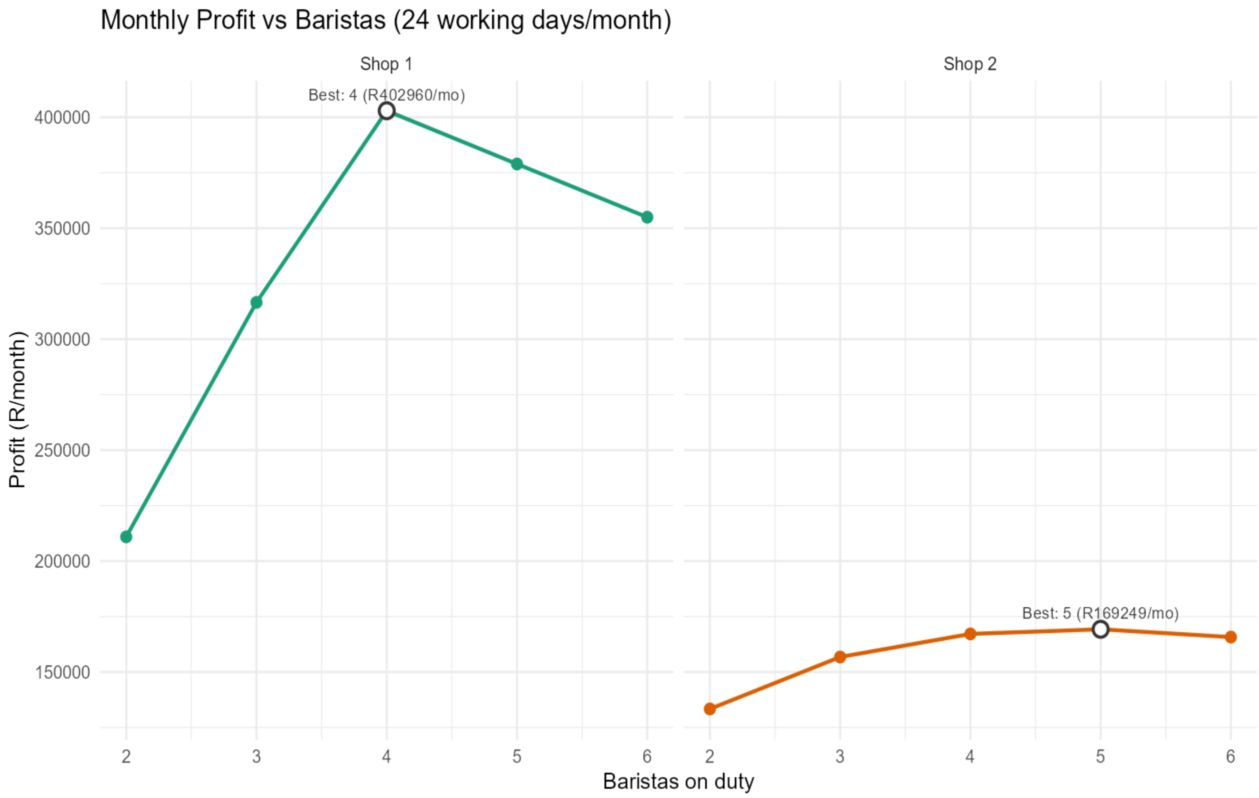

Figure 31: Monthly profit per number of baristas per shop

Figure 31 compares monthly profit across staffing levels. Shop 1's profit increases up to 4 baristas and then falls, indicating demand is fully met by B=4 (Profit = R402960/month) and additional staff are not revenue accretive. Shop 2's optimum is B=5 (R169249/month); further staffing yields negligible revenue benefit and reduces profit via payroll.

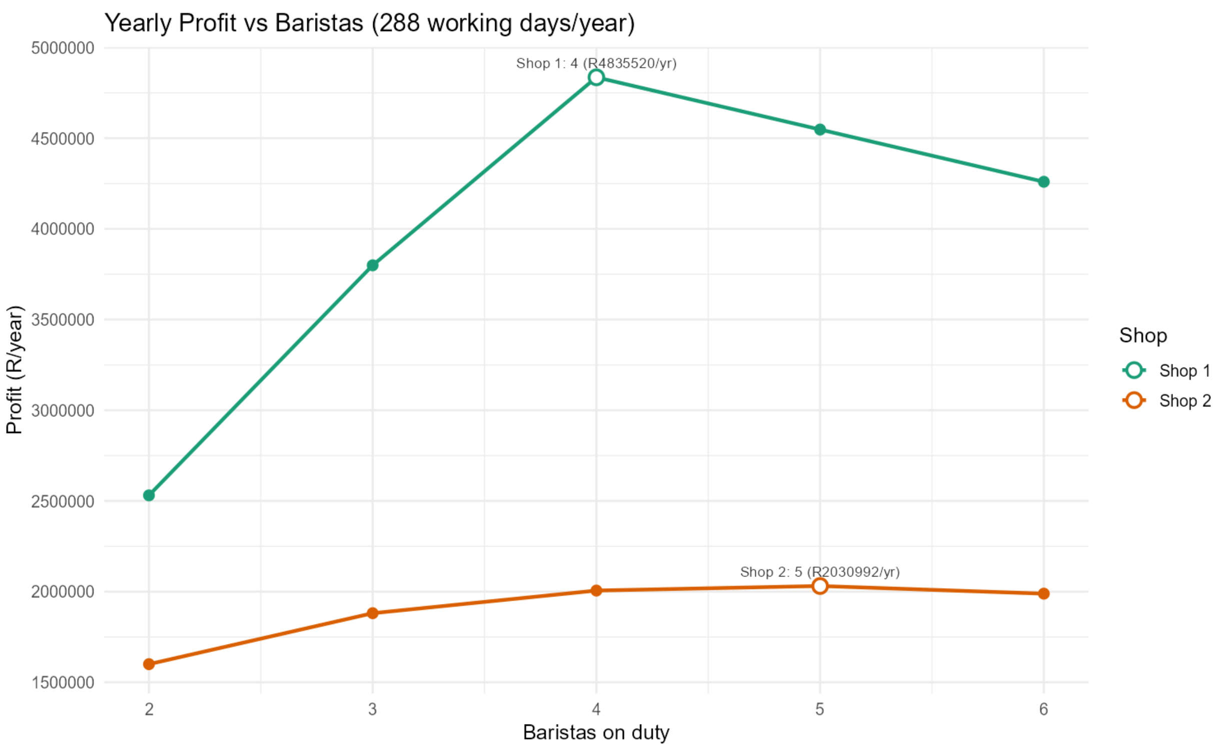

Figure 32: Yearly profit per number of baristas per shop

Figure 32 combines to yearly profit. Optima remain at 4 and 5 baristas for shops 1 and 2 respectively, confirming robustness over the working year. Shop 1's yearly profit with 4 baristas is R4835520 and shop 2's with 5 baristas is R2030992.

baristas	mean_sec	p95_sec	reliable_pct_under_T	shop
2	100.17098	112	99.7187852	Shop 1
3	66.61174	77	100.0000000	Shop 1
4	49.98038	59	100.0000000	Shop 1
5	39.96183	48	100.0000000	Shop 1
6	33.35565	41	100.0000000	Shop 1
2	141.51462	154	0.1241675	Shop 2
3	115.44091	126	79.3454067	Shop 2
4	100.01527	109	99.9971663	Shop 2
5	89.43597	98	100.0000000	Shop 2
6	81.64272	89	100.0000000	Shop 2

Table 3: Service reliability of staffing for shop 1 and shop 2

Table 3 summarises service performance by the number of baristas. In both shops, increasing staffing reduces both the average service time and the 95th-percentile time, indicating faster and more consistent service. Shop 1 operates substantially faster at every staffing level; with only two baristas, 99.7% of customers are served within 120 seconds, and reliability reaches ~100% from three baristas onwards. In contrast, shop 2 is capacity-constrained at low staffing: with two baristas virtually no customers meet the 120-second target (0.12%), improving to ~79% at three baristas and ~100% at four or more. The 95th-percentile in Shop 2 falls from 154 seconds at two baristas to 89 seconds at six baristas, confirming large gains in tail performance as staff increase.

baristas	n	mean_sec	p95_sec	reliable_pct_under_T	capacity_per_day	customers_served_day	revenue_day_R
2	3556	100.17098	112	99.71879	359.5693	359.5693	10787.08
3	12126	66.61174	77	100.00000	539.6964	539.6964	16190.89
4	29305	49.98038	59	100.00000	720.0523	693.0000	20790.00
5	56701	39.96183	48	100.00000	900.6374	693.0000	20790.00
6	97895	33.35565	41	100.00000	1081.4523	693.0000	20790.00

Table 4: Throughput of staffing for shop 1

baristas	n	mean_sec	p95_sec	reliable_pct_under_T	capacity_per_day	customers_served_day	revenue_day_R
2	8859	141.51462	154	0.1241675	251.8538	251.8538	7555.615
3	19768	115.44091	126	79.3454067	317.7042	317.7042	9531.126
4	35289	100.01527	109	99.9971663	365.4845	365.4845	10964.534
5	54958	89.43597	98	100.0000000	401.7352	401.7352	12052.056
6	78930	81.64272	89	100.0000000	430.1803	430.1803	12905.408

Table 5: Throughput of staffing for shop 2

Tables 4 (Shop 1) and 5 (Shop 2) summarise performance and capacity for each staffing level. In both shops, increasing baristas reduces the mean and 95th-percentile service times and raises daily service capacity. For Shop 1, reliability at 120 seconds is effectively 100% from B=3, and capacity exceeds demand from B=4, after which customers served per day and revenue saturate (~693 customers; ~R20,790/day).. Additional baristas mainly improve already-fast service and do not increase revenue. For Shop 2, low staffing results in poor reliability (0.12% ≤120 seconds at B=2) and insufficient capacity; reliability improves to ~100% at B≥4. Capacity meets demand at B≈5–6, after which revenue also saturates (~R12,905/day).

Conclusion:

Shop 1 is demand-limited at B≥4; thus 4 baristas maximise profit while maintaining excellent service. Shop 2 becomes demand-limited at B=5; 5 baristas achieves both full throughput and 100% reliability. Lower staffing in Shop 2 causes lost sales and slower service; higher staffing adds cost with little incremental revenue.

Question 6

6.2) DOE and ANOVA

Data selection

The same sales process data and response is used as in Question 3 (delivery time in hours, cleaned/ordered by timestamp). Two years are present in the file, 2022 and 2023, so those are analysed and also tested for a Month effect and a Year×Month interaction.

Hypotheses

- Year effect (one-way ANOVA)
  - $H_0: \mu_{2022} = \mu_{2023}$for delivery time.
  - H_a: At least one year mean differs.
- Month & interaction (two-way ANOVA)
  - H_0: No Month effect; H_0: no Year×Month interaction.

Methods

- Response: delivery_hours.
- Factors: Year (2 levels), Month (1–12).
- We fit a one-way ANOVA (Year) and, when Month was available, a two-way ANOVA (Year, Month, Year×Month).
- To visualise separability, a LDA was ran with equal class priors on standardised predictors.

Results

- Year (one-way ANOVA): $F(1,99,998) = 1.39$, p = 0.238.
  Effect size: $\eta^2 \approx 0.00001$, Cohen's d ≈ 0.008.

As seen in figure 33, there is no practical year-to-year difference in delivery time. Both years rise together through the year. There is a strong month affect, but no interaction.

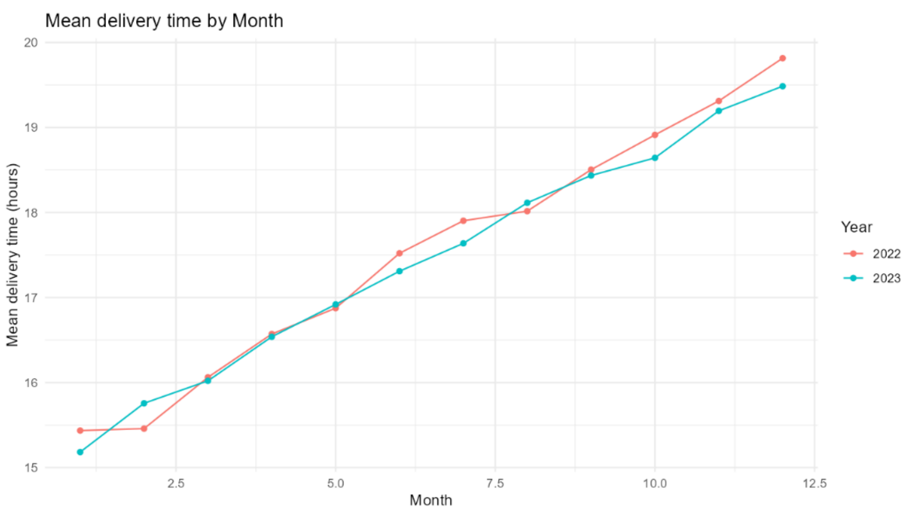

Figure 33: ANOVA mean delivery time by month

- Month (two-way ANOVA): Month is highly significant ($p < 2\times10^{-16}$). Year×Month interaction: $p = 0.742$.

Strong seasonality and the months later in the year have slower delivery times. The seasonal pattern is similar across years. Distributions overlapping almost perfectly supports a non-significant year effect.

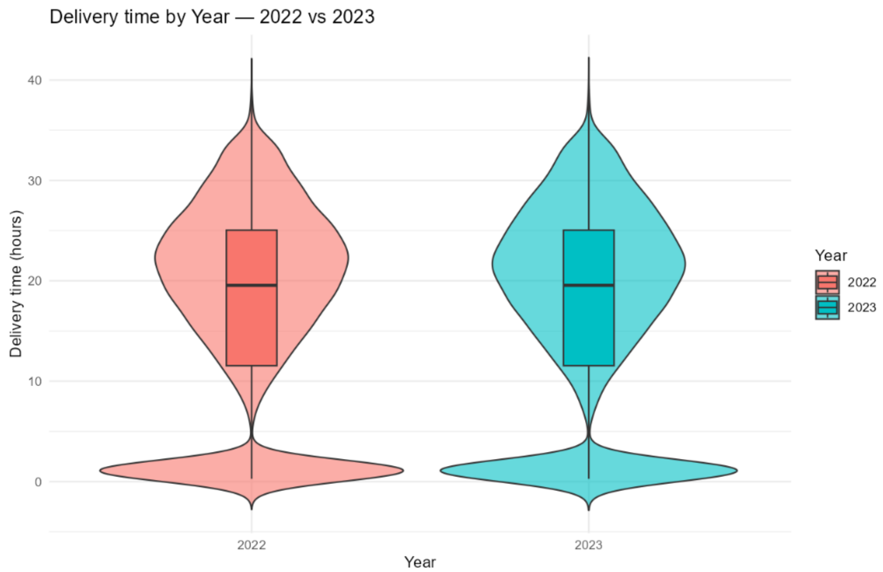

Figure 34: Violin/box plot of delivery times by year

- LDA check (2022 vs 2023): accuracy ≈ 50% (resub) and ≈ 50% (CV).

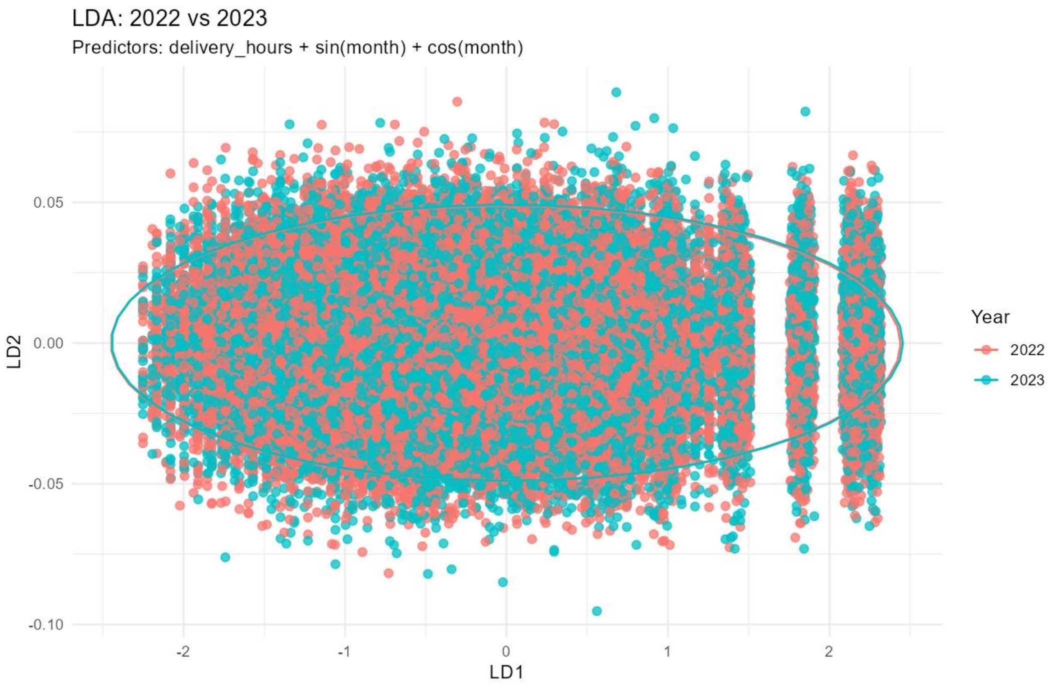

Figure 35: LDA scatter plot

This graph confirms that year labels are not separable using these features. Red and blue clouds overlap, and ellipses coincide meaning that there is no separability by year.

family	n_year1	n_year2	n_total	anova_p	eta2	cohens_d	lda_resub_acc	lda_cv_acc
CLO	8466	7132	15598	0.859595183	2.006339e-06	-0.002843144	0.5001282	0.4991666
KEY	9553	8367	17920	0.004505027	4.501926e-04	0.042535933	0.5105469	0.5105469
LAP	5446	4761	10207	0.481233241	4.861142e-05	-0.013974857	0.5033800	0.4944646
MON	8027	6837	14864	0.503576748	3.010354e-05	0.011008101	0.5067949	0.5051130
MOU	11108	9554	20662	0.466776117	2.563425e-05	0.010154455	0.4998064	0.4925467
SOF	11127	9622	20749	0.672327287	8.622712e-06	-0.005888143	0.5060485	0.5016627

Table 6: Per product family year test outcome table

The year analysis was repeated within each product family (CLO, KEY, LAP, MON, MOU, SOF). Sample sizes per family ranged from 10k to 20k observations across the two years. ANOVA p-values were mostly non-significant; where a small p-value occurred (e.g., KEY: p=0.0045), the effect sizes were trivial ($\eta^2 \leq 0.00045$; Cohen's d ≈ 0.04). Linear Discriminant Analysis (equal priors, standardized) achieved ~50% cross-validated accuracy for all families, i.e., no separability of years based on delivery time (and seasonal transforms). Conclude that there is no practically meaningful year-to-year shift in delivery times within any family. Combined with the aggregate two-way ANOVA, Month remains the dominant driver (seasonality), and the Year×Month pattern is stable.

Assumptions

ANOVA assumes independent observations, normal residuals, and equal variances. With $n = 100{,}000$, the CLT makes inference robust to mild non-normality. Thus, the conclusions are consistent.

Conclusion

There is no meaningful year-to-year shift in delivery time between 2022 and 2023. The dominant driver is the month (seasonality), and the seasonal pattern is stable across years. Operational planning should focus on seasonal capacity adjustments rather than year effects.

Question 7

7.1) Days per year of reliable service

h	problem_days_obs	reliable_rate	expected_reliable_days_year	expected_problem_days_year	expected_loss_year	hiring_cost_year
0	31	0.9219144	336.5	28.5	570025	0
1	6	0.9848866	359.5	5.5	110327	300000
2	1	0.9974811	364.1	0.9	18388	600000
3	0	1.0000000	365.0	0.0	0	900000
4	0	1.0000000	365.0	0.0	0	1200000
5	0	1.0000000	365.0	0.0	0	1500000

Table 7: Outcomes per extra hire

From the 397-day history, 366 days met the reliability target (≥15 staff), so the observed reliability is 92.2% (366 out of 397 days). That projects to 336.5 reliable days per year (95% CI ≈ 89.1%–94.4%), leaving 28.5 "problem days" per year when staffing drops below the threshold.

7.2) Profit optimisation

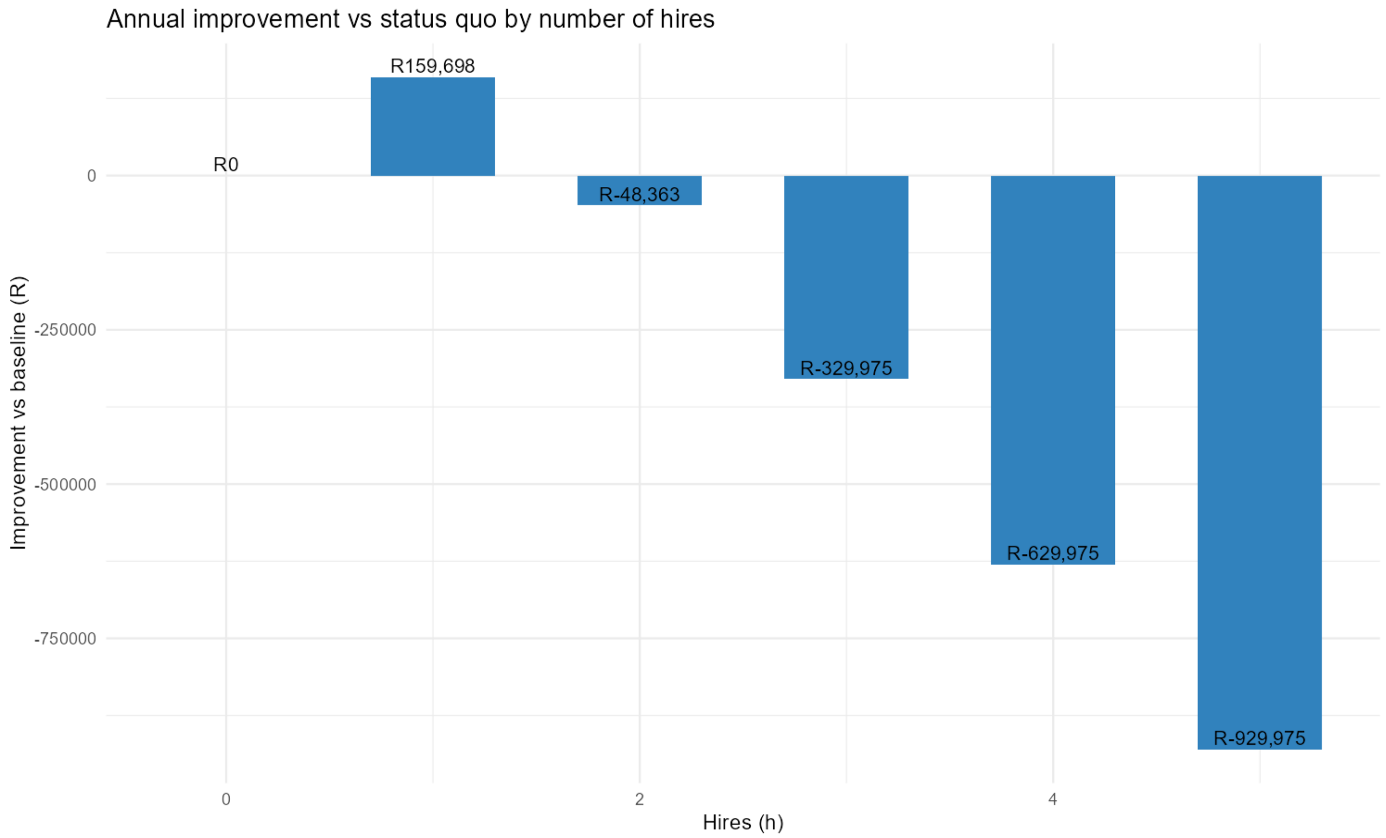

Figure 36: Annual improvement per number of hires

In figure 36 The bar at h=1 peaks above zero (label R159,698), showing the only improvement over the status quo. Bars at h≥2 drop progressively below zero, visualizing the diminishing returns and increasing salary burden once problem days are already scarce. Problem days reduce revenue by R20,000 day. Each hire costs R25,000 month = R300,000 year. The break-even condition for one hire is a reduction of at least 15 problem days per year (R300,000 / R20,000). Our estimates show a reduction of 23 days with the first hire (28.5 → 5.5), which exceeds break-even. A second hire only saves 4.6 additional days, which does not cover the extra R300,000 per year.

Recommendation: Hire exactly one additional staff member to optimise profit. This lifts reliability from ≈92% → ≈99% and adds R160k per year after costs. Hiring beyond one over-spends for negligible reliability gains, turning the decision net negative.

Conclusion

The project provided useful insights across all sections. The descriptive analysis showed that customers across different cities had similar income and spending habits, with only small differences. It also showed that product type matters more than price when it comes to profit—some product categories consistently performed better. The control charts revealed that delivery time variation stayed stable, but the average delivery time often shifted upward over time. These shifts suggest there may be underlying problems, such as policy or system changes, that need attention.

The process capability results showed that not all delivery processes meet the quality target (32 hours), either because the process is too variable or not centered correctly. The error analysis confirmed that false alarms (Type I errors) were rare, and sustained good control was evident in some product types. However, the risk of missing a real problem (Type II error) was quite high in certain cases, especially when the process changed slightly—this shows the need for better detection methods, like increasing sample size or using more rules.

The ANOVA results found no real difference in delivery performance between 2022 and 2023, but did show strong seasonal patterns, with slower deliveries later in the year. These trends were consistent across years, suggesting the need to plan resources based on monthly demand rather than yearly changes. In the staffing optimisation, adding baristas greatly improved service times and reliability up to a point—after that, the extra cost wasn't worth it. For example, shop 1 worked best with 4 baristas, and shop 2 with 5. A similar analysis for a car rental company showed that hiring one extra person saved the company money by reducing lost sales, but hiring a second wasn't cost-effective.

In summary, this project showed how statistical tools can be used to understand and improve real processes. The combination of descriptive statistics, control charts, capability analysis, and staffing models helped identify where performance could be improved. These methods allow engineers to make better, data-based decisions and highlight the importance of using evidence to guide operational and quality improvements in any system.

References

American Society for Quality (2025) ‘What is Statistical Process Control (SPC)?’, *ASQ Quality Tools*. Available at: https://asq.org/quality-resources/statistical-process-control (Accessed: 8 October 2025).

Feldman, K. (2024) ‘Process Capability Index: The Key to Customer Satisfaction and Business Success’, *iSixSigma*. Available at: https://www.isixsigma.com/dictionary/process-capability-index/ (Accessed: 8 October 2025).

Frady, L. (2024) ‘Process Capability (Cp, Cpk) and Process Performance (Pp, Ppk): What’s the Difference?’, *iSixSigma*. Available at: https://www.isixsigma.com/capability-indices-process-capability/process-capability-cp-cpk-and-process-performance-pp-ppk-what-difference/ (Accessed: 8 October 2025).

Qualtrics (n.d.) ‘ANOVA: definition, types, examples, and guide’. Available at: https://www.qualtrics.com/en-gb/experience-management/research/anova/ (Accessed: 16 October 2025).

Vanli, O. A. and del Castillo, E. (2021) ‘Statistical Process Control in Manufacturing’, *Encyclopaedia of Systems and Control*, pp. 2142–2150. Available at: https://doi.org/10.1007/978-3-030-44184-5_258 (Accessed: 15 October 2025).