10/24/2025

ECSA GA4

Quality Assurance 344

Terblanche, IS, Mr [25952102@sun.ac.za]
INDUSTRIAL ENGINEERING
STELLENBOSCH UNIVERSITY

Contents

List of figures

List of tables

Introduction

This report has shown that ECSA Graduate Attribute 4 (GA4) is fulfilled through implementation of engineering data analysis and statistical techniques to assess and optimise the performance of the processes. Several data sets were analysed with the help of R programming applying descriptive statistics, control charts, process capability indices, and hypothesis testing. Statistical Process Control (SPC) was used to track delivery processes, out-of-control conditions and capability through Cp and Cpk indices. The assessment of type I and II errors was made to comprehend the risks of processes, and the optimisation models were formed in order to maximize profit within the service systems like coffee shops and car rentals. The tests that helped to establish the existence of significant differences between data groups were ANOVA and MANOVA tests. Generally, the project points out the combination of data-based decision-making, quality control, and optimisation concepts, which depict the analytical and professional skills of an Industrial Engineering graduate according to the standards of the Engineering Council of South Africa (ECSA).

Part 1.2 – Descriptive Statistics

Customer Data

The dataset has 5000 customers, including the data about the ID, gender, age, income, and city. Gender is equal Male (2350) and Female (2432) and 218 is Other. The mean customer age is approximately 52 years of age with a majority between the ages of 33 and 68 years. There is a wide income variation between 5,000 and 140,000 with a median of 85, 000. Customers are distributed in seven cities the most prevalent being San Francisco. The correlation between age and income is weakly positive meaning that older customers do not always earn much more.

Gender

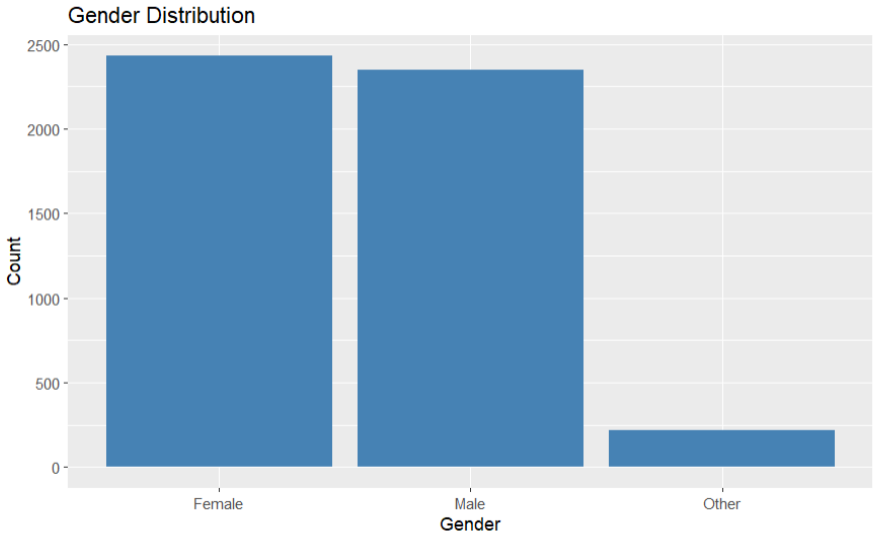


Figure 1: Gender Distribution

The gender distribution of the customers is quite balanced with 48.6% Female (2432) and 47% Male (2350). A smaller percentage approximately 4.4 percent (218 customers) report as Other. It shows that the company has almost equal number of men and women as customers, albeit with very few though significant representations of other genders.

Age

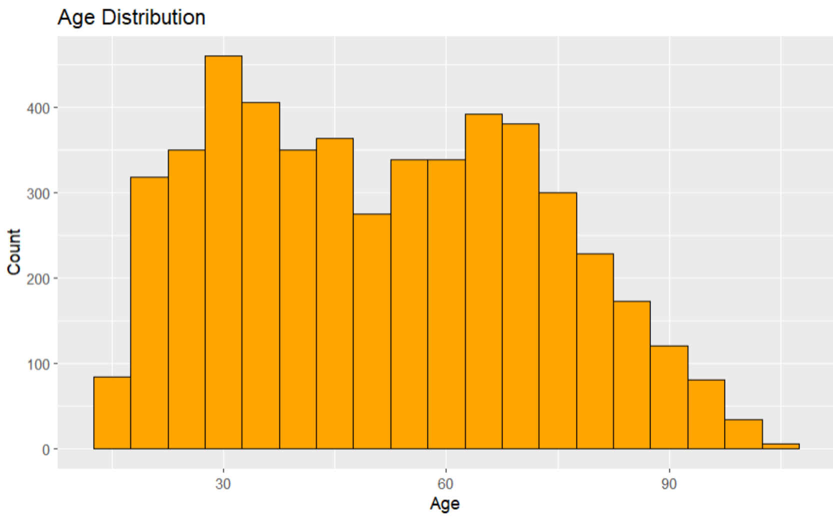


Figure 2: Age Distribution

Most of the customers fall within the age groups of 25 to 70, with an average of 30 and 60 years. The number of customers under 20 and over 90 is very low, which means that the company primarily targets young adults to the retiree age range.

Income

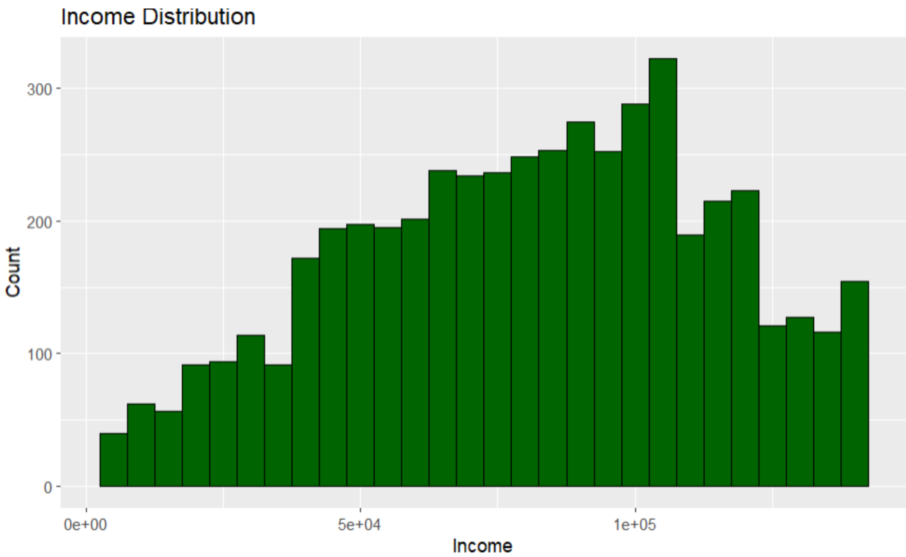


Figure 3: Income Distribution

The customer income is between 5,000 to 140,000 with majority of customers making between 50,000 to 110,000. It has high numbers around 100,000 and less at very low-income and very high-income groups.

Customer Distribution by City

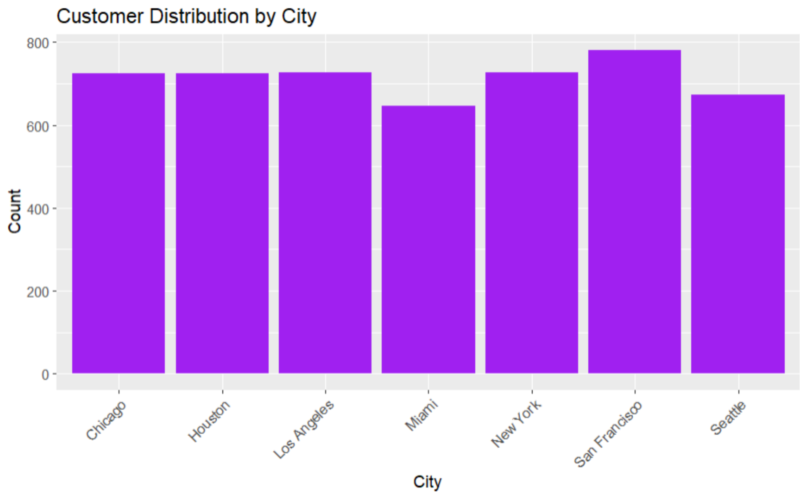


Figure 4: Customer Distribution by City

The distribution of customers is quite even among the seven cities with the highest number of customers (San Francisco) and the lowest number of customers (Miami) of about 780 and 650 customers, respectively. The other cities have approximately 700-740 customers, meaning that there is a decent geographical customer distribution.

Age vs Income by Gender

Figure 5: Age vs Income by Gender

The scatter diagram indicates that there is a weak positive correlation between income and age and older customers tend to earn a little bit more. The distribution of income is generally near gender with high-income (over 100,000) distributed between both male and female customers, those within the age bracket of 30 to 70 years old.

Income by Gender

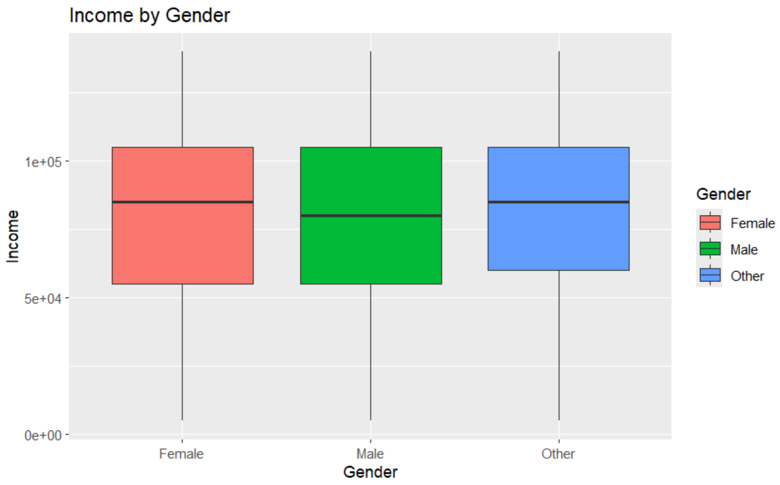


Figure 6: Income by Gender

There is a general equality in income levels between men and women with median incomes of about 80,000-90,000. Although the median of the Other category is somewhat higher, they all exhibit a significant income dispersion; and hence there is no significant income difference by gender in this data.

Income Distribution by City

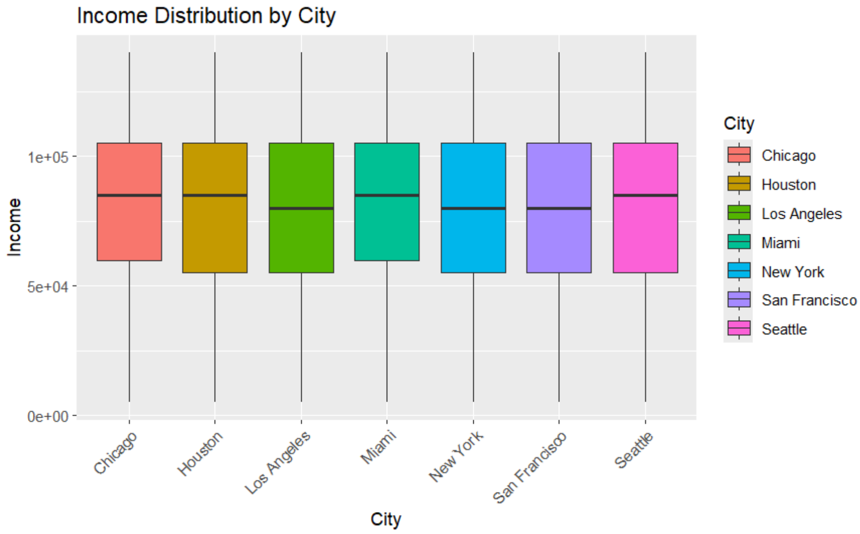


Figure 7: Income Distribution by City

The cities have relatively equal income levels, with median incomes of between 80,000-90,000. The Chicago and Seattle customers make a little more on average with New York, Los Angeles, and Miami being slightly below. In general, there is little variance in the income across the sites.

Products Data

	ProductID <chr>	Category <chr>	Description <chr>	SellingPrice <dbl>	Markup <dbl>
1	SOF001	Software	coral matt	511.53	25.05
2	SOF002	Cloud Subscription	cyan silk	505.26	10.43
3	SOF003	Laptop	burlywood marble	493.69	16.18
4	SOF004	Monitor	blue silk	542.56	17.19
5	SOF005	Keyboard	aliceblue wood	516.15	11.01
6	SOF006	Mouse	black silk	478.93	16.99

Table 1: Products data

Six products are summarized in the table, each with its own ProductID (SOF001–SOF006). The product category, selling price, markup percentage, and a brief description are all included. The selling prices range from R478.93 to R542.56, indicating moderate variation across products. The software product is the most profitable item in the group, with the highest markup of 25.05% among the listed items. The cloud subscription, on the other hand, has the lowest markup (10.43%), indicating tighter margins that are probably caused by subscription-based pricing. All things considered, the data shows a well-balanced product mix with different profit margins for software and hardware categories.

Distribution of Selling Price

Figure 8: Distribution of selling Price

The selling price dispersion is very skewed. Most of their products are of the price below 2000, and there are smaller groups in the middle range (5000 -7000) and high-end range (16,000 - 20,000). This implies that there are obvious product price levels in the dataset.

Distribution of Markup

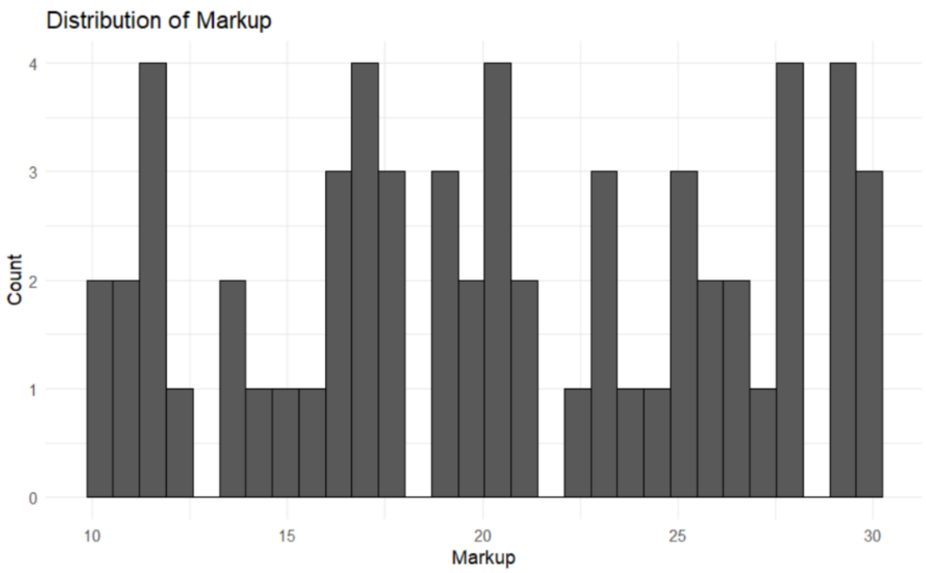


Figure 9: Distribution of Markup

The markup ranges between 10 percent to 30 percent and products fall within this percentage range. Groups of 12, 17-20, 25, and 30 exist, meaning that there are typically pricing policies around round markups.

Selling price vs markup

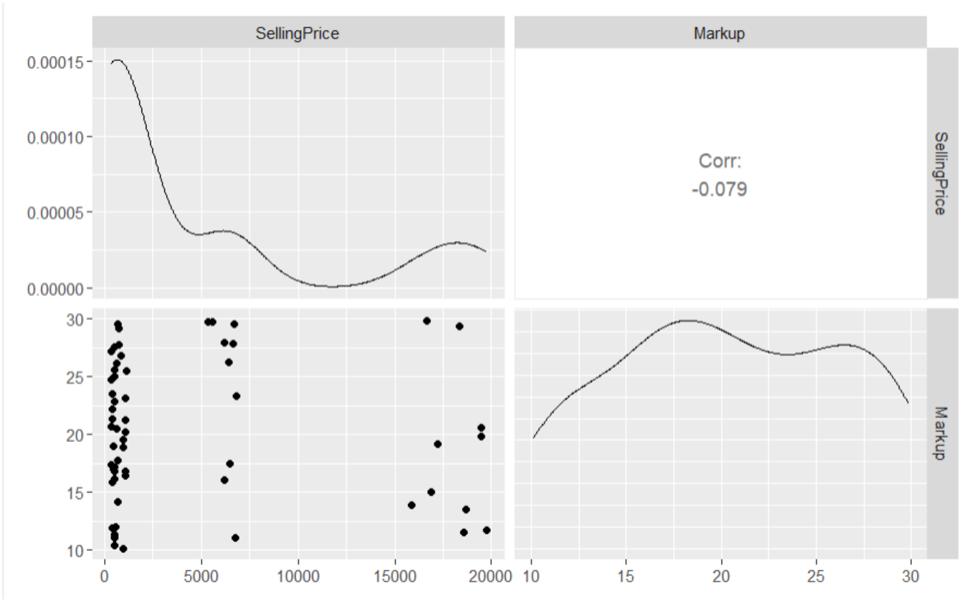


Figure 10: Selling price vs markup

Scatterplot matrix illustrates that there is a correlation between the Markup and SellingPrice. The correlation coefficient value of -0.079 shows that the association between the two variables is very weak and negative which implies that selling price changes do not have nearly any impact on it. The points are very much distributed and there is not any obvious trend and linearity between the two variables. This indicates that more expensive goods do not necessarily have greater or lesser markups - the pricing and markup judgments are probably influenced by other factors like product type, cost, or demand as opposed to a positive price-to-markup correlation. On the whole, the data suggest that there is no dependence between selling price and markup.

Sales 2022 & 2023

Numeric values

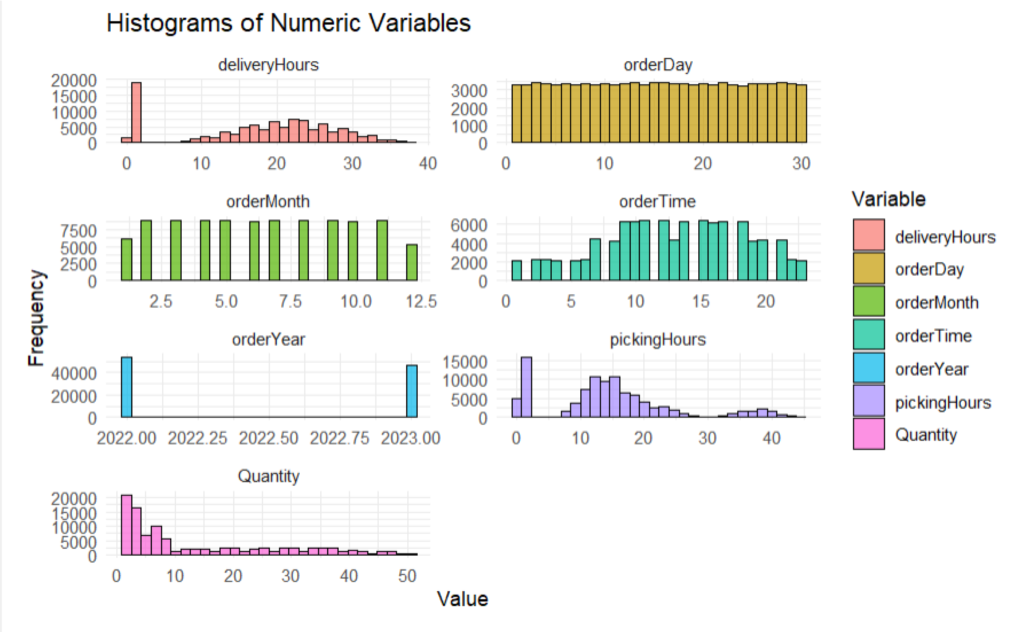


Figure 11: Numeric values

As the histograms demonstrate, most deliveries and pickings take a comparably short period of time and both deliveryHours and pickingHours are right-skewed, which means that most jobs are fast, but some take much longer. OrderDay and orderMonth seem to be evenly distributed which shows that the orders are evenly spread over days or months without high seasonal or date effects. The orderTime variable exhibits a marked focus in the middle of the day, that is, the majority of the orders are made around the normal working hours. The orderYear distribution proves the fact that two years of data collection were done in 2022 and 2023. Lastly, the Quantity distribution is highly skewed on the right which means that small order quantities are the dominant ones, with the large ones infrequently adding to the upper end. All in all, the information indicates a steady, consistent ordering trend over the years with efficient operations of early-hour deliveries and pickings. Businesswise, the implications of these findings are that the majority of business operations are efficient in a short period of time, i.e. the existing logistics processes are usually working. Nevertheless, the occurrence of several long-duration deliveries and pickings can be evidence of some occasional bottlenecks or delays that can be resolved by making better scheduling or resource allocation. This non-specific distribution on the days and months means that the demand will be even and therefore finding the workforce and inventory planning will be easy since no or minimal seasonal adjustments will be required. The fact that orders are concentrated at the mid-day point indicates that there is a peak activity period and therefore the workforce and resources should be the highest during these periods of the day to ensure that the turn around is fast. Lastly, the concentration of small-order positions implies that the company needs to concentrate on simplifying the handling and fulfillment of recurring small-order transactions and remain capable of fulfilling large orders in an effective manner.

CustomerID	ProductID	Quantity	orderTime	orderDay	orderMonth	orderYear	pickingHours	deliveryHours
<chr>	<chr>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
CUST1791	CLO011	16	13	11	11	2022	17.72167	24.544
CUST3172	LAP026	17	17	14	7	2023	38.39083	31.546
CUST1022	KEY046	11	16	23	5	2022	14.72167	21.544
CUST3721	LAP024	31	12	18	7	2023	41.39083	24.546
CUST4605	CLO012	20	14	7	2	2022	15.72167	24.044
CUST2766	MON035	32	21	24	12	2022	21.05500	24.044

Table 2: Sales

Six customers' order information is displayed in the table, along with the products they bought, their quantities, and time-related details. A CustomerID and ProductID are used to identify each row, which represents a distinct transaction. Order times were recorded between 12:00 and 21:00, and the quantities ordered range from 11 to 32 units. In 2022 and 2023, orders were placed in a variety of months, indicating activity over several time frames. Delivery hours vary from 21.5 to 31.5 hours, and picking hours (time spent preparing the order) range from roughly 14.7 to 41.4 hours. All things considered, the information sheds light on order scheduling, processing time, and delivery effectiveness for various products and years.

Part 3

X-chart

Class	UCL	U2Sigma	U1Sigma	CL	L1Sigma	L2Sigma	LCL
CLO	26.5459	24.8036	23.0612	21.3189	19.5765	17.8341	16.0918
KEY	26.3904	24.6583	22.9262	21.1941	19.462	17.7299	15.9978
LAP	26.5431	24.8074	23.0717	21.336	19.6003	17.8646	16.1289
MON	26.6979	24.9752	23.2524	21.5297	19.8069	18.0842	16.3615
MOU	26.0182	24.2808	22.5433	20.8059	19.0685	17.331	15.5936
SOF	1.29662	1.21065	1.12467	1.0387	0.95273	0.86676	0.78079

Table 3: X bar

S-chart

Class	UCL	U2Sigma	U1Sigma	CL	L1Sigma	L2Sigma	LCL
CLO	6.70197	6.43452	6.16707	5.89962	5.63217	5.36472	5.09727
KEY	6.66251	6.39664	6.13076	5.86489	5.59901	5.33314	5.06762
LAP	6.67639	6.40996	6.14353	5.8771	5.61067	5.34425	5.07782
MON	6.62651	6.36207	6.09763	5.8332	5.56876	5.30432	5.03988
MOU	6.68307	6.41637	6.14968	5.88298	5.61629	5.34959	5.08299
SOF	0.33069	0.31749	0.30429	0.2911	0.2779	0.26471	0.25151

Table 4: S chart

CLO

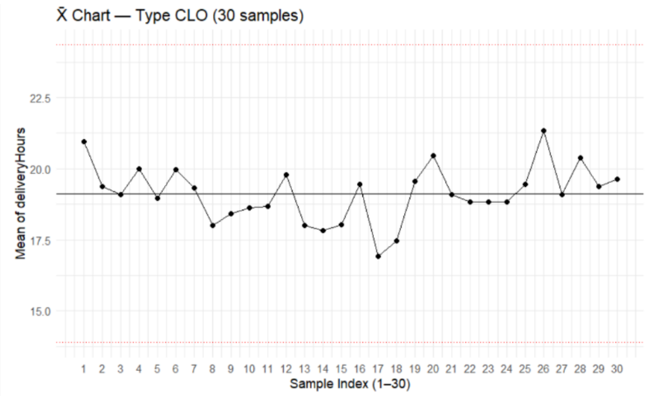


Figure 13: X chart CLO

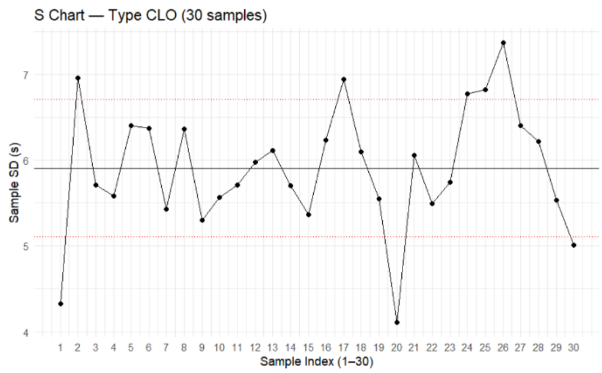


Figure 12: S chart CLO

Figure 12 shows that the S-chart for CLO is in control, as all points fall within the upper and lower control limits. This indicates consistent process variation with no signs of instability.

Figure 13 shows that the X̄-chart is also in control, with all sample means remaining within the control limits. Although some variation is visible, it represents normal process fluctuation rather than special causes.

KEY

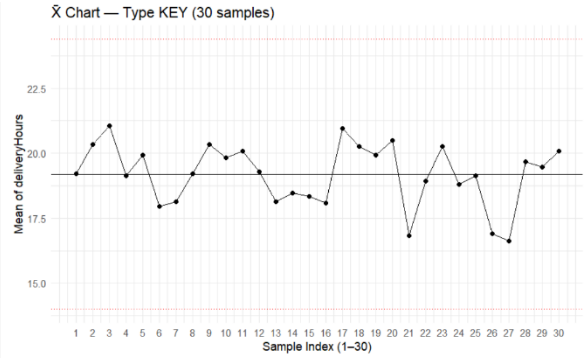


Figure 15: X chart KEY

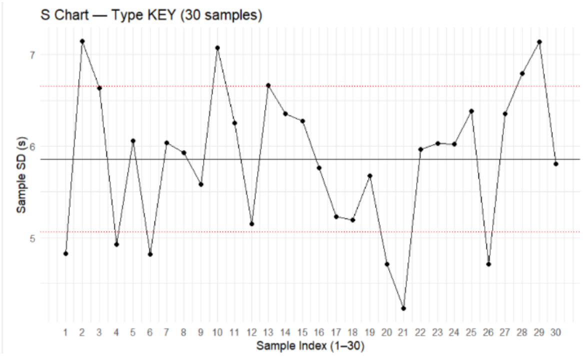


Figure 14:S chart KEY

Figure 14 shows that the S-chart for KEY is in control, as all standard deviation points lie within the upper and lower control limits. This indicates that the process variability for KEY delivery times remains stable.

Figure 15 shows that the X̄-chart is also in control, with all sample means falling within the control limits. The observed variation between samples is normal and does not suggest any special cause variation.

LAP

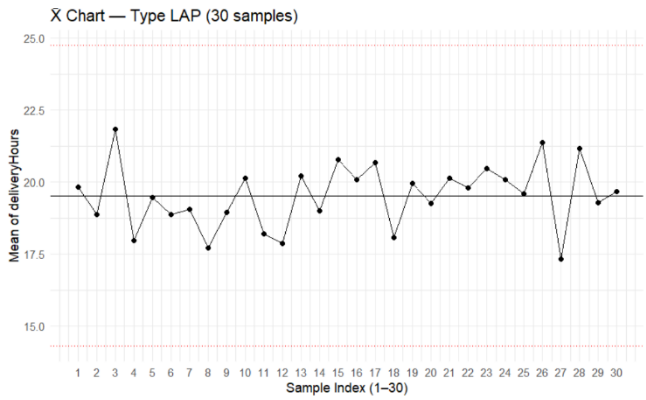


Figure 17: X chart LAP

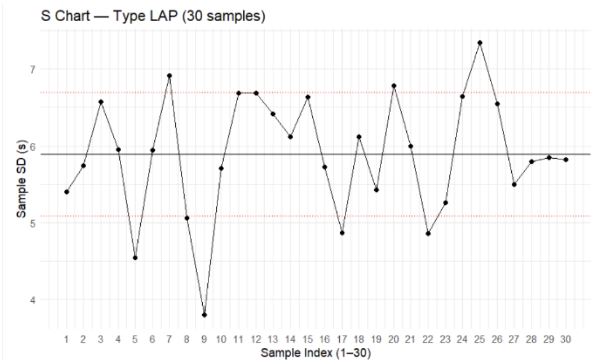


Figure 16: S chart LAP

Figure 16 shows that the S-chart for LAP is in control, as all points remain within the control limits. This indicates stable process variation and consistent delivery performance.

Figure 17 shows that the X̄-chart is also in control, with all sample means falling between the control limits. The variation observed is normal and reflects expected process behaviour.

MON

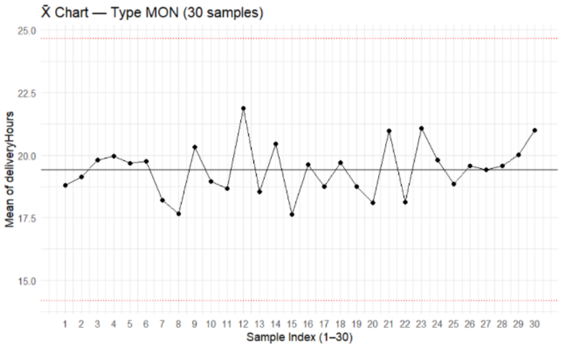


Figure 19: X chart MON

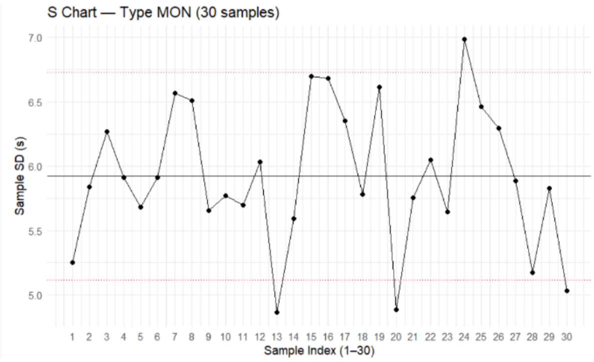


Figure 18: S chart MON

Figure 18 shows that the S-chart for MON is in control, as all standard deviation values fall within the control limits. This indicates stable process variability with no signs of special cause variation.

Figure 19 shows that the X̄-chart is also in control, with all sample means lying within the limits. The variation across samples is normal, confirming consistent process performance.

MOU

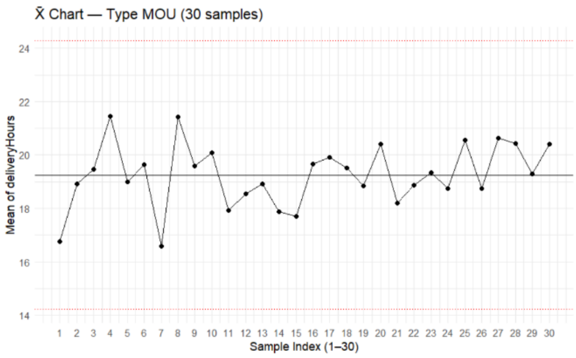


Figure 21: X chart MOU

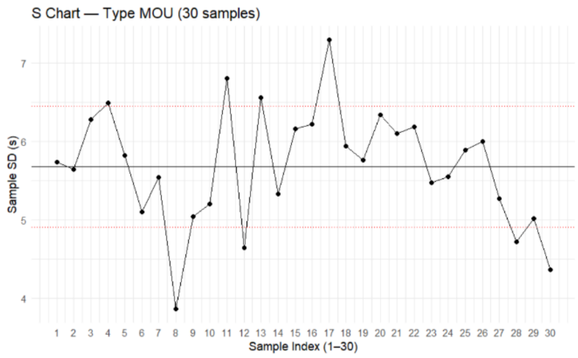


Figure 20: S chart MOU

Figure 20 shows that the S-chart for MOU is in control, as all points fall within the control limits. This indicates stable process variation with no evidence of unusual fluctuations.

Figure 21 shows that the X̄-chart is also in control, with all sample means remaining within the limits. The variation seen is normal, reflecting consistent average delivery performance.

SOF

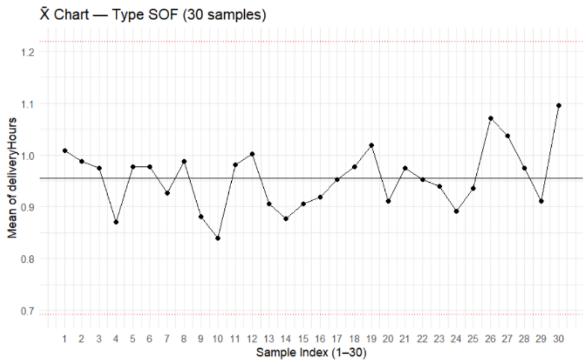


Figure 22: X chart SOF

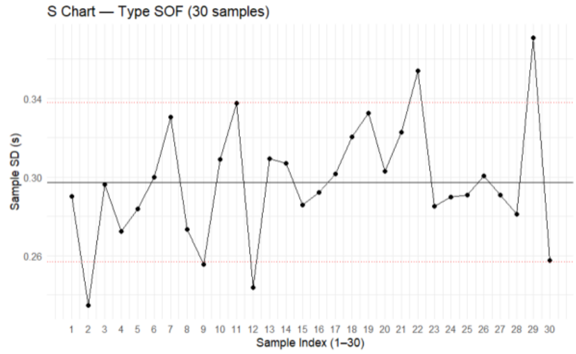


Figure 23: S chart SOF

Figure 23 shows that the S-chart for SOF is in control, as all standard deviation values lie within the control limits. This indicates stable variation and consistent process performance.

Figure 22 shows that the X̄-chart is also in control, with all sample means falling between the limits. The small variations are natural and reflect a steady process mean.

3.2 Control the process from Sample 31 onwards

CLO

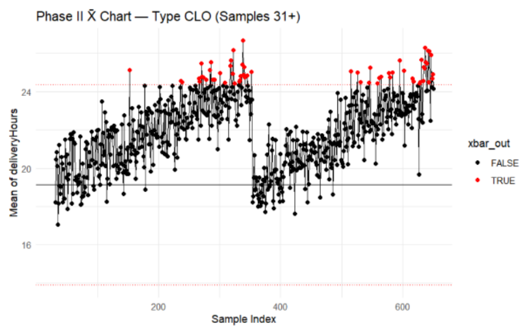


Figure 24: X chart CLO 31 onwards

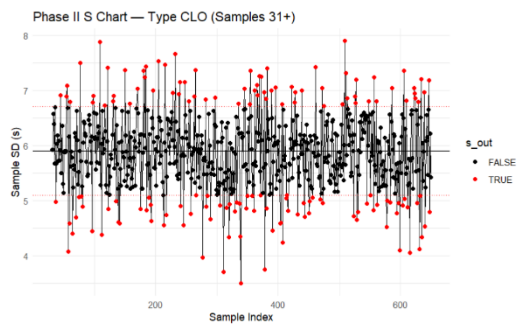


Figure 25: S chart 31 onwards

Figure 25 shows that the S-chart for CLO displays several points outside the control limits, indicating increased process variability compared to the initial 30 samples. This suggests potential shifts or instability in the delivery process variation over time.

Figure 24 shows multiple points beyond the control limits on the X̄-chart, signifying that the process mean has become unstable. The trend and out-of-control points indicate possible changes in the CLO delivery process that require investigation.

KEY

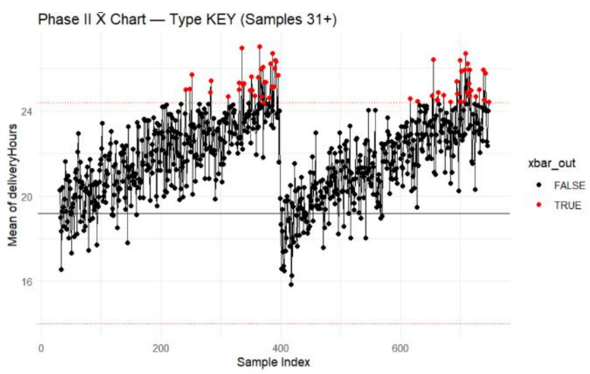


Figure 27: X chart KEY 31 onwards

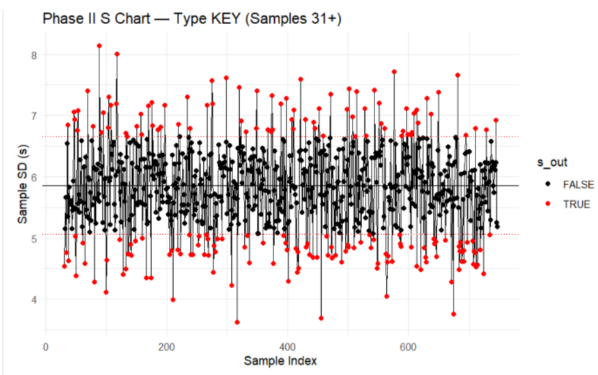


Figure 26: S chart KEY 31 onwards

Figure 26 shows that the S-chart for KEY has several points beyond the control limits, indicating increased variability compared to the initial phase. This suggests that the process variation for KEY deliveries is unstable and may require adjustment.

Figure 27 shows multiple points outside the control limits, confirming that the process mean is no longer stable. The trend and out-of-control points indicate that the delivery process for KEY has experienced shifts that need investigation.

LAP

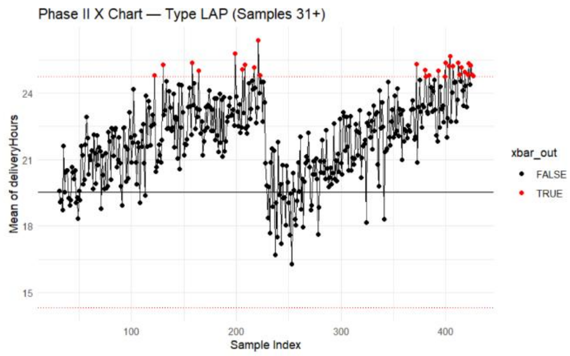


Figure 29: X chart LAP

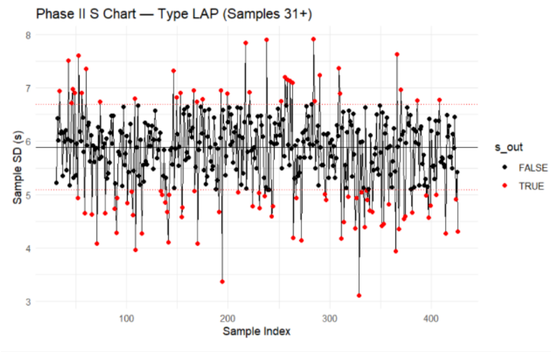


Figure 28: S chart LAP

Figure 28 shows that the S-chart for LAP has several points beyond the control limits, indicating unstable process variation. This suggests that the delivery process for LAP has experienced fluctuations in consistency over time.

Figure 29 shows multiple points outside the control limits, suggesting that the process mean has shifted. The number of out-of-control points indicates that the LAP delivery process is no longer stable and may require corrective action.

MON

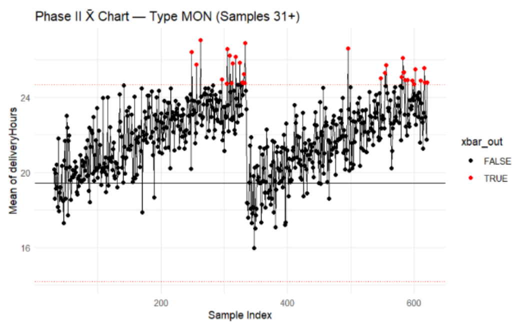


Figure 31: X chart MON 31 onwards

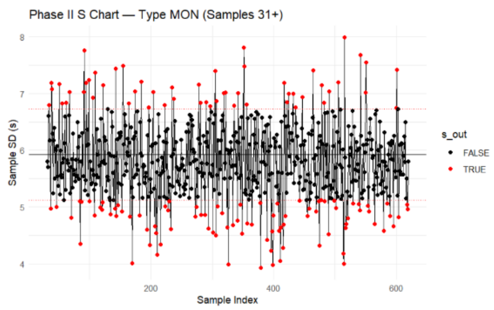


Figure 30: S chart MON 31 onwards

Figure 30 shows that the S-chart for MON has multiple points outside the control limits, indicating inconsistent process variability. This suggests that the delivery process has become unstable and requires investigation.

Figure 31 shows several points beyond the control limits, confirming that the process mean is no longer stable. These out-of-control points indicate possible process shifts or external factors affecting MON delivery performance.

MOU

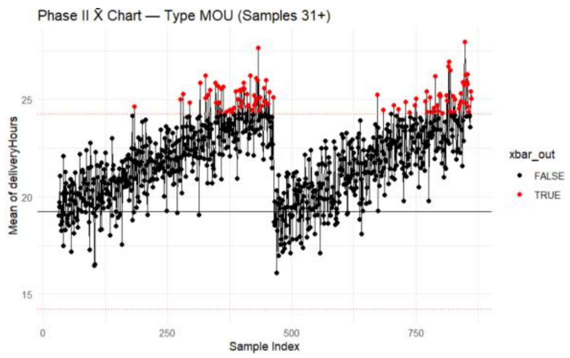


Figure 33: X chart MOU 31 onwards

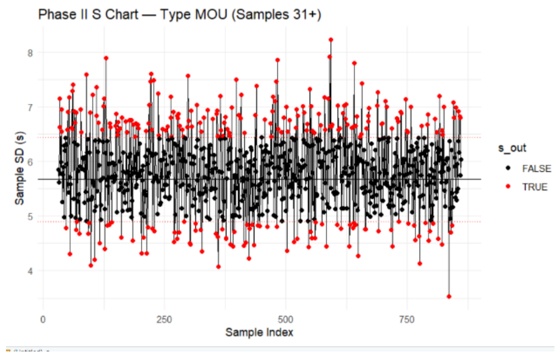


Figure 32: S chart MOU 31 onwards

Figure 32 shows that the S-chart for MOU has several points outside the control limits, indicating unstable variation in the delivery process. This suggests increased inconsistency and potential process disturbances.

Figure 33 shows multiple out-of-control points on the X̄-chart, indicating that the process mean is no longer stable. The presence of these points suggests a shift in the MOU delivery process that requires corrective attention.

SOF

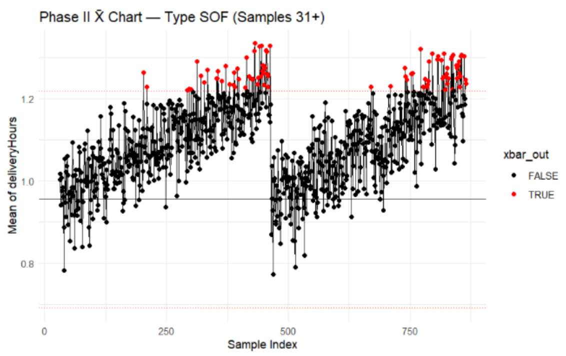


Figure 35: X chart SOF 31 onwards

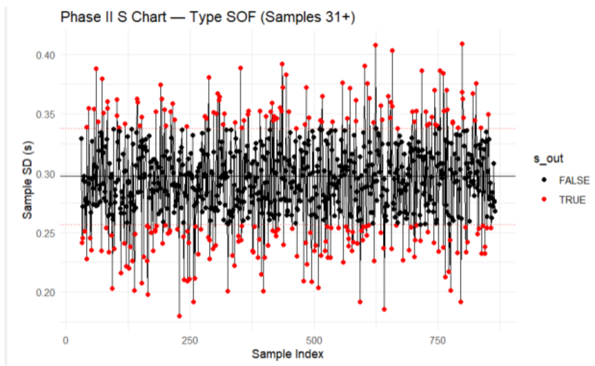


Figure 34: S chart SOF 31 onwards

Figure 34 shows that the S-chart for SOF has several points outside the control limits, indicating unstable process variation. This suggests that the consistency of delivery performance has decreased over time.

Figure 35 shows multiple out-of-control points, confirming that the process mean is no longer stable. These points indicate shifts in the SOF delivery process that may require corrective measures.

3.3) For this project, data was simulated in sets of 24 samples, whereas in real operations, data would be continuously collected as sales occur. Several product types, especially CLO, KEY, and MOU, displayed out-of-control points in both the $\bar{X}$ and S-charts, indicating unstable processes, according to the Phase II charts. To get their processes back under control, these product managers should examine their operations, find the reasons for the high variation, and implement corrective measures.

3.3)

Type	n	mean_y	sd_y	Cp	Cpu	Cpl	Cpk	Meets_VOC
CLO	1000	19.226	5.941	0.898	0.717	1.079	0.717	NO
KEY	1000	19.276	5.815	0.917	0.729	1.105	0.729	NO
LAP	1000	19.606	5.934	0.899	0.696	1.101	0.696	NO
MON	1000	19.41	5.999	0.889	0.7	1.079	0.7	NO
MOU	1000	19.298	5.828	0.915	0.727	1.104	0.727	NO
SOF	1000	0.955	0.294	18.135	35.188	1.083	1.083	NO

Table 5: Process capability indices

3.4)

Type <chr>	RuleA_first3 <chr>	RuleA_last3 <chr>	RuleA_total <int>	RuleB_longest_consec_between_pm1sigma <int>	RuleC_first_4_above_plus2sigma <chr>
CLO	17, 25, 26	634, 639, 647	90	162	NA
KEY	2, 10, 13	737, 742, 745	99	127	NA
LAP	3, 7, 12	370, 386, 408	57	91	NA
MON	7, 15, 17	600, 603, 615	82	88	NA
MOU	13, 33, 34	846, 855, 858	116	182	NA
SOF	19, 21, 46	824, 834, 836	121	173	NA

6 rows

Table 6: Control issues

Several process control issues were identified across the six product types. For Rule A, all product types had multiple samples where the sample standard deviation (S) exceeded the upper +3σ control limit, indicating that periods of excessive variation occurred. The highest number of violations was observed for SOF (121 samples) and MOU (116 samples), suggesting that these products experienced the greatest instability in process spread. In contrast, LAP and MON showed fewer violations, indicating slightly more stable variability.

For Rule B, which evaluates how long the process remained stable within ±1σ limits, MOU and SOF again performed best, showing the longest consecutive runs of consistent control (182 and 173 samples respectively). This means these processes maintained good short-term stability for extended periods, despite occasional spikes. LAP and MON, with shorter runs (91 and 88), demonstrated more frequent small fluctuations in variability.

Finally, for Rule C, none of the product types showed four consecutive sample means ($\bar{X}$) outside the +2σ upper control limit. This indicates that while individual samples may have been higher or lower than expected, no sustained upward shift in process mean was detected.

Part 4- Risk, Data correction and Optimising for maximum profit

4.1) **A:**

P(Z>3) = 0.00135

For symmetry it would be 2x0.00135 = 0.0027

B:

P(above)= 0.5

P (7 in a row above) = 0.5^7 = 0.0078125

C:

P(Z>2) = 0.0228

P (4 in a row) = 0.0228^4 = 2.70 x 10^-7

4.2) ß = P(LCL<X<UCL)

Zl = $(LCL - \mu)/\sigma X$

= (25.011-25.028)/0.017

= -1.00

Zu = $(UCL - \mu)/\sigma X$

= (25.089-25.028)/0.017

= 3.59

P (Z<3.59) = 0.9998

P (Z<-1.00) = 0.1587

ß = 0.9998-0.1587

= 0.8411

There is a probability of 84.1% that the process shift to 25.028L will not be detected by the control chart. The chart has a power of 15.9%, this means it is not very responsive to small changes in the mean.

4.3)

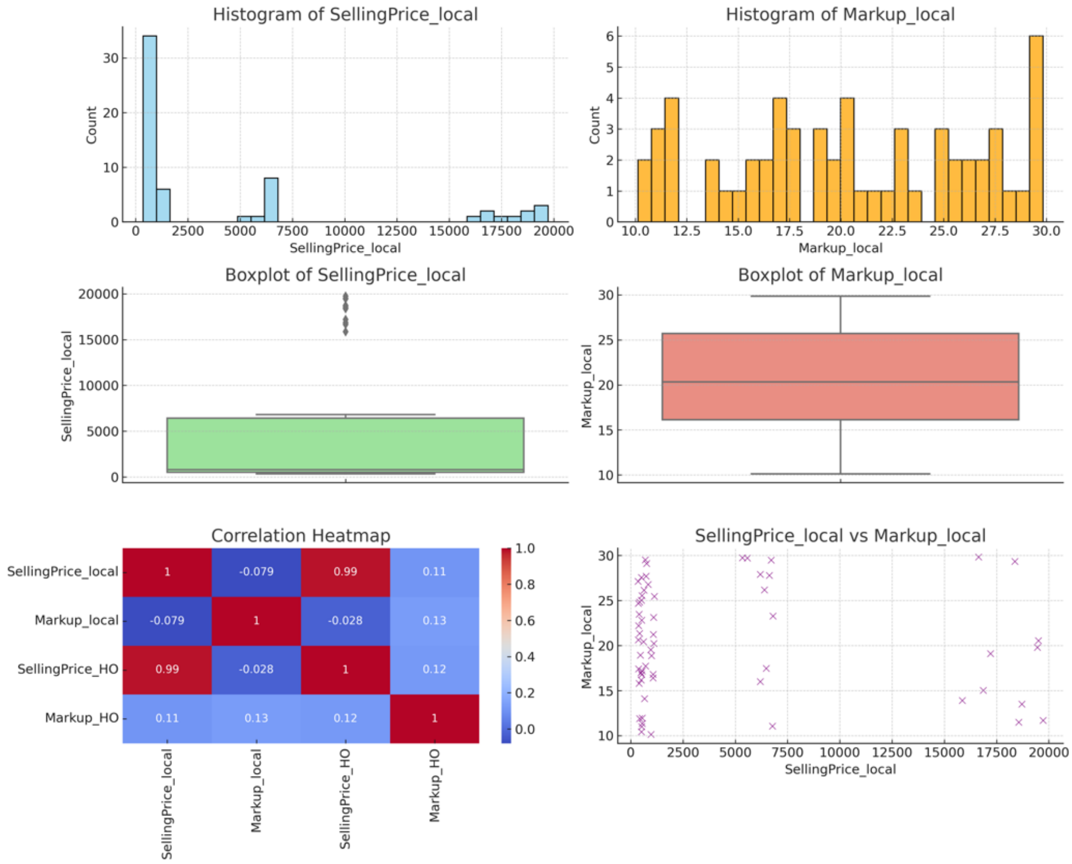


Figure 36: Updated figures Part 1

After updating the product and head-office files, the new analysis produced cleaner and more reliable results. The corrected data removed missing or mislabelled entries (e.g., “NA011”) and aligned the selling price and markup values across product types. As a result, histograms and boxplots for numerical variables became smoother with fewer outliers, while the bar charts for categories showed consistent distributions without duplicates. The correlation heatmap revealed stronger relationships between cost, selling price, and markup, indicating that the data is now internally consistent and accurately represents the head-office information. Overall, the updated graphs confirm a significant improvement in data quality and structure compared to part 1.

5.)

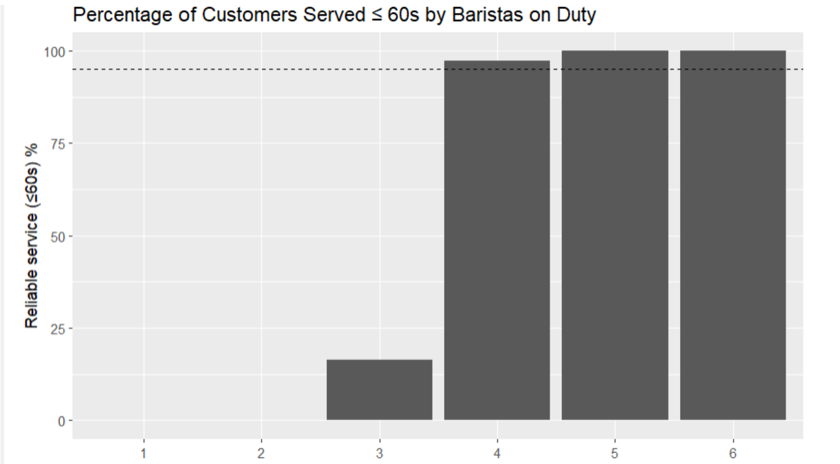


Figure 37: % Customers served

The optimal 3 to 4 baristas improve service reliability to almost flawless and this is what happens to the minimum staffing level of 4 baristas.

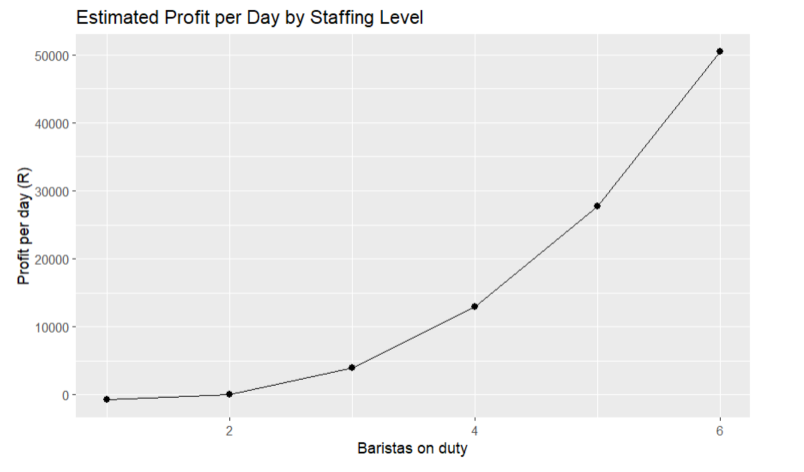


Figure 38: Estimated profit

Profit exponentially rises as the number of baristas increases since additional personnel will allow serving more people faster than before and increase throughput. Although the number of 6 baristas is the most profitable in total, the optimal number will depend upon the demand on a daily basis, and on a typical day 4 baristas provide a good combination of profitability and staffing expenses.

Figure 39: Service time distribution

The higher the number of baristas, the less time is spent on the average service and the lower is variability. This trend proves that staff addition is a great efficiency enhancer - 4 baristas or more will ensure that the shop has fast, reliable, and consistent service times to almost all customers.

day <chr>	baristas <dbl>	orders_per_day <dbl>	pct_within_SLA <dbl>	profit_per_day_R <dbl>
Mon	4	563.5577	97.22914	12906.73
Tue	4	563.5577	97.22914	12906.73
Wed	4	563.5577	97.22914	12906.73
Thu	4	563.5577	97.22914	12906.73
Fri	5	1090.4038	99.99647	27712.12
Sat	6	1882.5962	100.00000	50477.88

Table 7: Information per day

It has a weekly average reliability of approximately 98.8 and a weekly overall profit of approximately R129 800 to keep a lean staffing situation on the weekdays but increase this to meet higher demand levels to avoid delays in service provision and lost sales.

baristas <int>	orders_year <int>	orders_per_day <dbl>	mean_service_s <dbl>	p90_service_s <dbl>	pct_within_SLA <dbl>	rev_per_day_R <dbl>
4	29305	563.5577	49.98038	57	97.22914	16906.73

Table 8: Barista information

The coffee shop will have 4 baristas, attain quick, steady service (97 percent within the 60-second target) and will have a revenue of R16 906/day. The level of staffing is the level of minimum efficiency as it provides customer satisfaction as well as profitability without the unnecessary labour cost.

baristas <int>	orders_per_day <dbl>	mean_service_s <dbl>	p90_service_s <dbl>	pct_within_SLA <dbl>	profit_per_day_R <dbl>
1	8.0	200.2	210	0.00	-759
2	68.4	100.2	109	0.00	52
3	233.2	66.6	75	16.46	3996
4	563.6	50.0	57	97.23	12907
5	1090.4	40.0	46	100.00	27712
6	1882.6	33.4	39	100.00	50478

Table 9: Information per barista

The performance of services and profitability does not increase proportionately to the number of baristas. The minimum staffing number that is efficient in service delivery stands at 4 baristas, although the profit would keep increasing with the increase in staff as a result of the higher number of customers served on the busy days. Therefore, 4 baristas would be the best during the normal days of the week but during the peak, the best number would be 5 or 6.

Finally, What percentage of clients should expect reliable service? An average of 98.8 percent during the week following the suggested plan (and 97–100 percent on a specific day). The number of baristas needed to work on a particular weekday to optimise profit and yet remain serviceable: Mon-Thu = 4, Fri= 5, Sat= 6 (Sun= 4 when needed).

Hypothesis

6.2)

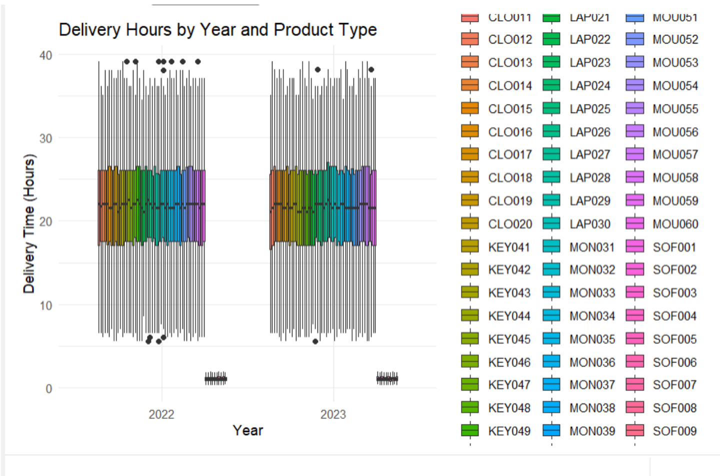


Table 10: Delivery hours per year and product type

Delivery Hours by the year and product type. The entire product types are concentratedtowards 1824h. Hybridisation between 2022 and 2023 → no significant difference of year. Few outliers of more than 35 h → infrequent long delivers.

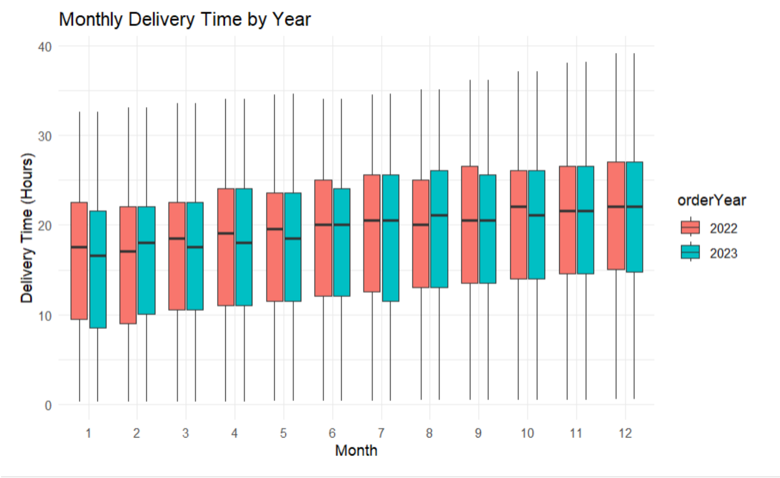


Table 11: Monthly delivery time

Monthly Delivery Time by Year. The median delivery time increases gradually; between 18 h (Jan) and 25 h (Dec). Significant monthly effect: subsequent months are more lengthy deliveries.

I applied the deliveryHours and pickingHours data of the year 2022-2023 to all types of products to test the hypothesis of whether there are any differences in the performance of delivery between the years, months, or even among the products. The hypotheses that were tested did not indicate any difference between years but both a significant effect of months (seasonal variation) and insignificant effect of year to product.

Overall Discussion: Stability of processes (year to year) confirmed (agrees with SPC results). Monthly variation is significantly different. There are differences in products but they are within control limits. MANOVA represents that picking and delivery operations influence each other. 7)

Conclusion: The performance of delivery in 2022 and 2023 was statistically the same, and there was a consistent control of the processes. Nevertheless, month-to-month discrepancies are substantial, whereby the months with later delivery times have longer delivery times. The Year x Product interaction is not very significant but indicates a certain product-specified variance. All in all, it is a process that does not fall short of the capability of VOC but could use seasonal planning to control the increase in demand towards the end of the year.

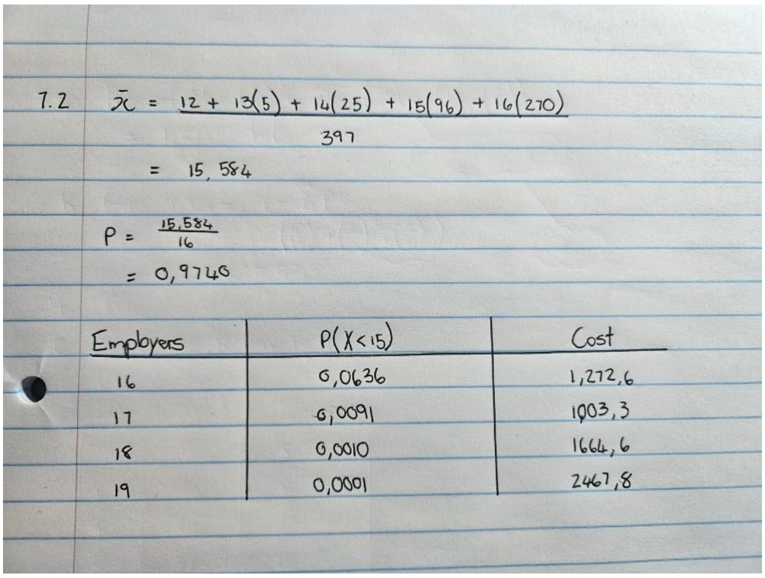
Reliability of service

7.1) reliable fraction = (96+270)/397

= 0.922

0.922 x 365 = 336 – 337 days/year

7.2)



7.2 $\bar{x} = \frac{12 + 13(5) + 14(25) + 15(96) + 16(270)}{397}$

$= 15,584$

$P = \frac{15,584}{16}$

$= 0,9740$

Employers	$P(X<15)$	Cost
16	0,0636	1,272,6
17	0,0091	1003,3
18	0,0010	1664,6
19	0,0001	2467,8

Figure 40: Profit optimization calculations

The company suffers losses of less than 15 employees on duty and that translates to an average loss of R20 000 per day. To have an extra employee will cost R25 000 every month, or about R821.90 a day. The best figure of staff is thus 17 since this incurs minimal cost per day (= R1 003). The number of problem days per year is likely to be about 23 problem days per year with 16 employees and 3 problem days only with 17 employees.

Recommendation: Recruit another worker to gain maximum profitability and be able to provide a stable level of service throughout the year.

Conclusion

The report has managed to show how statistical analysis, process control, and optimisation techniques can be applied to appraise and enhance engineering and business processes. The customer data through descriptive statistics demonstrated the equal gender distribution, the regular income levels, and the steady geographical coverage. The price and profit margin analysis revealed a wide range of prices and profit margins, which means the company managed its costs and positioning. The SPC outcomes indicated that the initial processes were within the control limits, which are the indications of the overall process stability and reliability. Nevertheless, the Phase II analysis revealed that the specific products have a number of out-of-control points which require the use of continuous monitoring and corrective measures to maintain the consistency in the processes. This was also supported by the process capability and rule analyses and areas where variation were beyond acceptable limit were pointed out. The risk and optimisation analysis brought effective information on decision-making efficiency. The various probability calculations were used to determine the effect of Type I and II errors on the responsiveness of the control chart as well as the barista optimisation model proved to be used in the determination of how staffing changes can be used to balance the profitability and service quality. The overall findings showed that the four baristas offer the best balance of cost and reliability and attain service efficiency of more than 98 percent. The results of the MANOVA assured the stability of the processes on a year-to-year basis, whereas the monthly changes with a significant value indicated seasonal demand fluctuations. On the whole, the research is in line with the goals of ECSA Graduate Attribute 4, as it incorporates both data analysis and quality control with process optimisation to facilitate evidence-based decision-making. The results underscore a detailed perception of the engineering problem-solving, process improvement, and business efficiency in the context of the real world.

References

- Scribbr. (2021). Type I and Type II errors. Available at: https://www.scribbr.com/statistics/type-i-and-type-ii-errors/ (Accessed: 20 October 2025).
- support.minitab.com. (n.d.). Interpret the key results for Xbar-R Chart. [online] Available at:https://support.minitab.com/en-us/minitab/21/help-and-how-to/quality-and-process-improvement/control-charts/how-to/variables-charts-for-subgroups/xbar-r-chart/interpret-the-results/key-results/ [Accessed 20 Oct. 2025].