ECSA REPORT

Shelby Nicholls

23923954

Nicholls, S, Miss [23923954@sun.ac.za]

Contents

Introduction

This report provides a comprehensive analysis for the Quality Assurance 344 2025 ECSA project (Engineering Counsel of South Africa, 2025). The analysis begins with a descriptive data analysis of the customer data and the product data, identifying key characteristics and initial data quality issues and a time-based analysis of sales patterns across 2022 and 2023 to understand temporal trends. Statistical Process Control (SPC) is used to assess the stability and capability of the product delivery processes. This analysis is supplemented by an evaluation of statistical risks, such as Type I and Type II errors, and the impact of data correction on product pricing analysis.

Furthermore, the report explores profit optimization models for two separate retail shops to determine ideal staffing levels. An Analysis of Variance (ANOVA) is used to statistically test the factors influencing delivery times. The report concludes with an assessment of service reliability, synthesizing the findings to provide a holistic view of operational efficiency and areas for strategic improvement.

Part 1: Descriptive Analytics

1.1 Customer Analysis

Insights from customer data:

- 5000 row entries and 5 column variables: CustomerID, Gender, Age, Income, City
- The minimum age of the customers is 16 with a maximum age of 105. The mean age of customers, 51.55, is very similar to the median, 51.
- The income ranges from 5000 to 140 000 with a mean of 80797 and a median of 85000
- There are 2432 females and 2350 males. There are 218 customers who chose other for their gender.
- There are no missing values from this dataset

As can be seen in Figure 1 below, most of the customers are between the ages of 30 to 35. The number of customers is also particularly high between the age of 25 to 30 and again between the ages of 65 to 70. The ages of the customers gradually increase from 16 to 35, stays relatively consistent until the spike from 65 to 70 and tapers off from 70 onwards. Majority of the customers are younger than 35 years old.

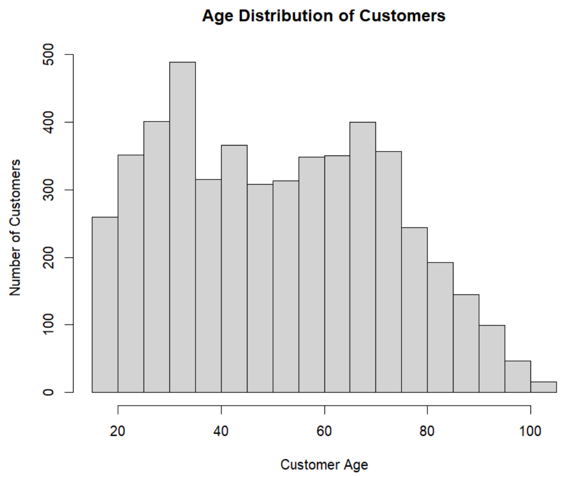


Figure 1: Age Distribution of Customers

The income distribution of customers indicates that most customers earn between 80 000 and 100 000 by the peak in the graph in Figure 2. The distribution drops off after 120 000, indicating fewer high-income customers. The customers are primarily middle-income customers with a smaller segment of high-income customers.

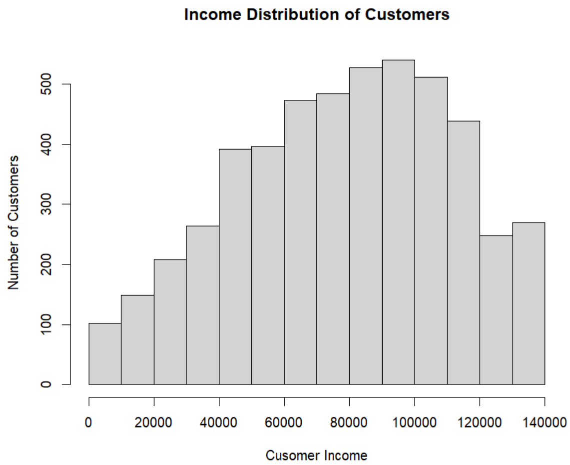


Figure 2: Income Distribution of Customers

As can be seen in Figure 3 and Figure 4 below, there is no significant correlation between the income of a customer and the city they live in nor the age of a customer and the city they live in.

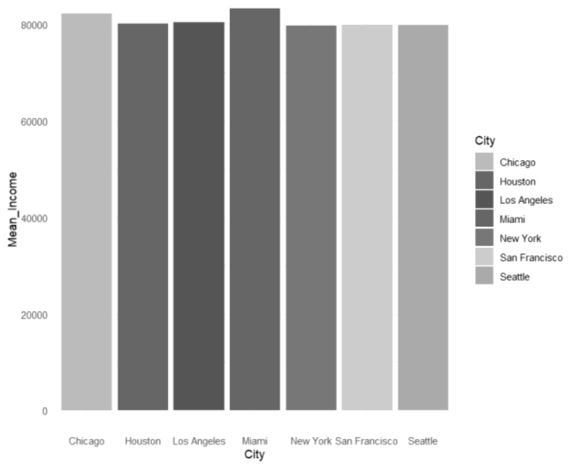


Figure 3: Mean Income by City

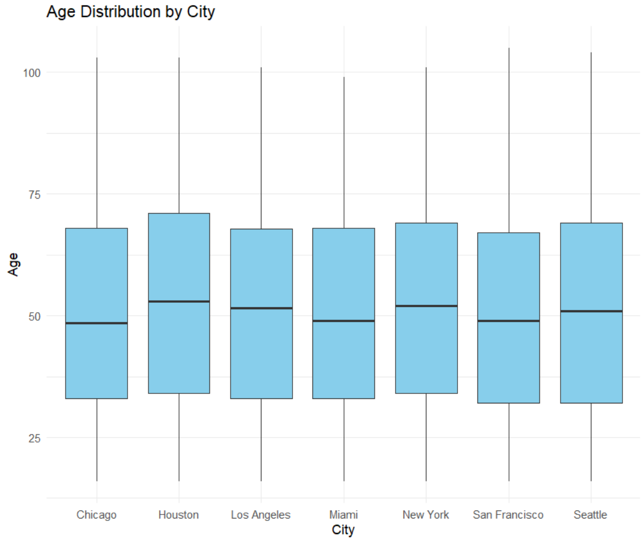


Figure 4: Age Distribution by City

Figure 5 indicates the relationship between the age of a customer and their income. The highest earning customers are between the ages of 40 to 50 and 50 to 60.

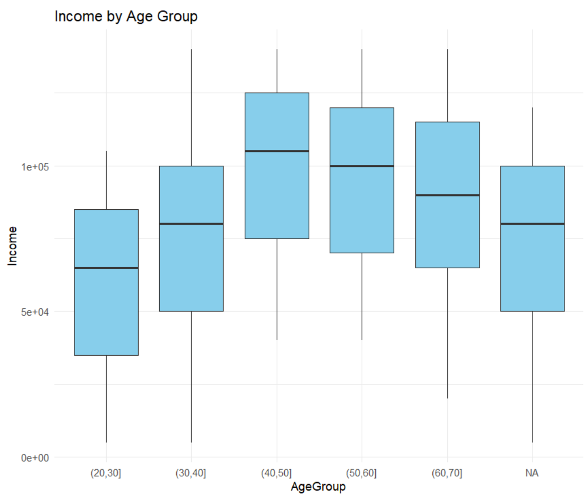


Figure 5: Income by Age Group

1.2 Products Data Analysis:

Only the first 10 products from the products_data are also in the products_Headoffice database. The products_Headoffice database has more entries but there are some products in products_data that are not in the head office database. This is concerning as it indicates that the local branch that the products_data was obtained from is selling items that the head office is not aware of. It is also concerning that the branch is not selling a big part of the head office product list. This could be a result of an outdated database. This inconsistency in the business can cause problems in the future when synthesizing data for analysis or reporting sales and other critical information to shareholders. This can also lead to inefficiencies like stockouts as the head office is unaware of the stock level of the unknown items at the local branch and therefore is unable to replenish them. The local branch is also potentially missing sales opportunities by not carrying the full product range from the head office. My recommendation to upper management is to remedy this situation to ensure all databases are updated and relevant. For the purpose of this basic data analysis, these two databases were merged and kept uniquely by the productID so that the analysis could proceed with all information for all products.

Insights from products data:

- 160 row entries and 5 column variables: ProductID, Category, Description, Selling Price and Markup.
- The company as a whole, the branch and the head office together, offer 160 unique products.
- The minimum selling price is 350,4 and the maximum is 20 348,4 with a mean of 3483,6 and a median of 545,5. The mean is greater than the median meaning that the data is right skewed so fewer products with a high selling price pulls the average selling price up.
- The minimum markup is 10,09 and the maximum is 29,94 with a mean of 20,49 and a median of 20,70. Since the mean and median markup are similar values, the company uses a consistent pricing strategy so there are no outlying products with an abnormally high or low markup.
- There are 6 product categories: Cloud Subscription, Keyboard, Laptop, Monitor, Mouse and Software. All product categories have 20 products each except for software which has 60 products. This indicates that the company's product range is predominantly software. This can be seen in Figure 6 below.
- There are no missing values.

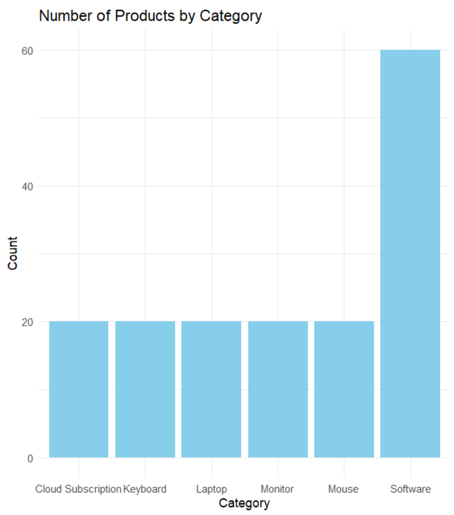


Figure 6: Number of Products by Category

Figure 7 shows the spread of selling price within each category. The selling price distribution is similar for cloud subscriptions, keyboard, laptop, monitor and mouse categories. These all have low price ranges compared to software with a few high-price outliers. The software distribution shows a large variation in selling price. This could be intentional to offer different tiers to customers, or it reflects a lack of pricing strategy.

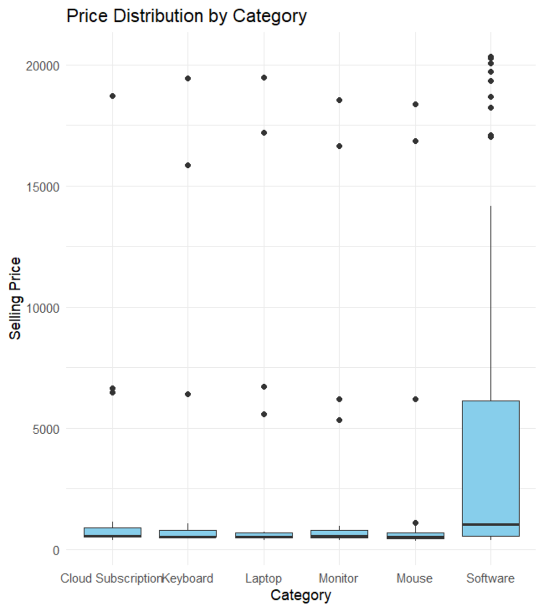


Figure 7: Price Distribution by Category

Most categories have a markup between 18% and 22% so the markup is fairly consistent across all categories. Cloud subscription and software have the highest markups out of all the categories.

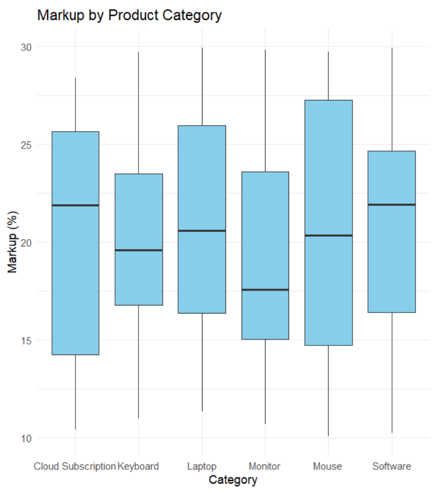


Figure 8: Markup by Product Category

The average profit was calculated by multiplying the selling price with the markup percentage in each category. The software category generates the most profit for the company. Laptops generate the second highest profit for the company, closely followed by mouse, monitor and keyboard. The cloud subscription category generates the least profit.

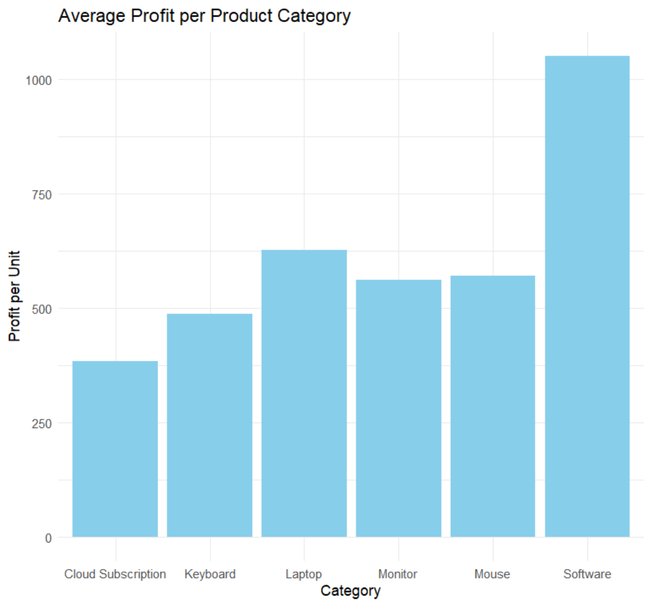


Figure 9: Average Profit per Product Category

The mean price is larger than the median price across all categories, indicating that there are very few expensive products pulling the average up.

Part 2: Time-Based Analysis

2.1 Sales Data Analysis

Insights from sales data:

- 100 000 row entries and 9 column variables: CustomerID, ProductID, Quantity, orderTime, orderDay, orderMonth, orderYear, pickingHours and deliveryHours.
- There were 53 727 sales made in 2022 and 46 273 in 2023.
- There are no missing values in this dataset.

Summary statistics for the sales 2022 and 2023:

Table 1: Sales Summary Statistics

	Quantity	orderTime	orderDay	orderMonth	orderYear	pickingHours	deliveryHours
Min	1	1	1	1	2022	0.4259	0.2772
Max	50	23	30	12	2023	45.0575	38.0460
Mean	13.5	12.93	15.5	6.448	2022	14.6955	17.4765
Median	6	13	15	6	2022	14.0550	19.5460

The number of sales decreased from 2022 to 2023 whilst still having similar sales trends. In both years, the monthly sales are always low in January and December. Potentially due to the business closing over the holiday period. In both years, the sales remain relatively consistent from February to November with a slight increase in sales from June to July.

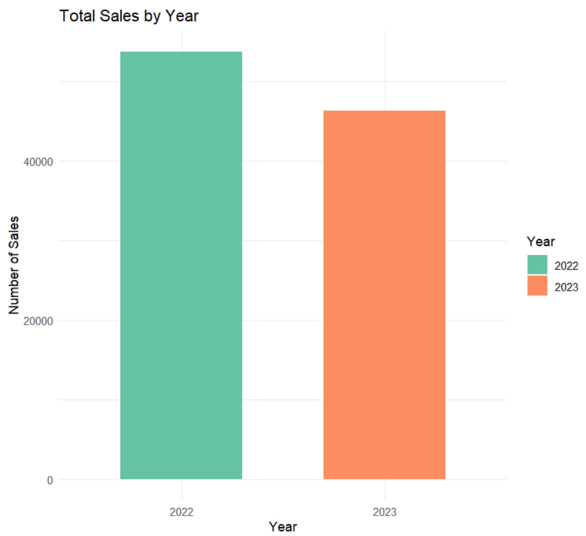


Figure 10: Total Sales by Year

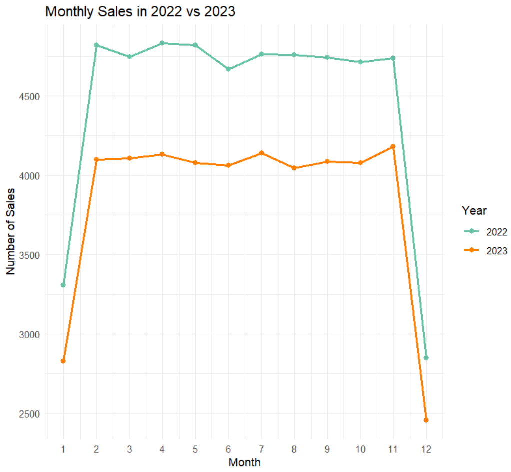


Figure 11: Monthly Sales Comparison

The average daily sales for 2022 and 2023 indicate a trend with spikes in sales on the 3rd, 13th and 22nd days of the month.

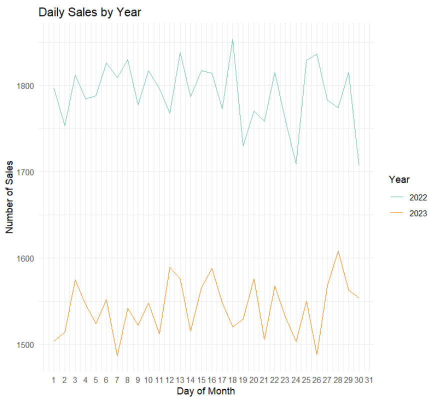


Figure 12: Daily Sales by Year

Sales peak from 9am to 12pm and again from 2pm to 6pm on average daily. Sales are lowest between 10pm and 6am. This trend mimics the daily 8-hour work day as customers make fewer sales because in the morning as they are getting ready and heading to work, eat lunch at 1pm and then go home from work after 6pm.

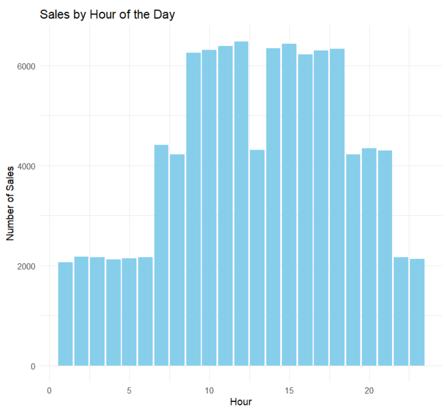


Figure 13: Sales by Hour of the Day

Most orders contain 5 products or less. Few customer orders have a quantity of 10 or larger.

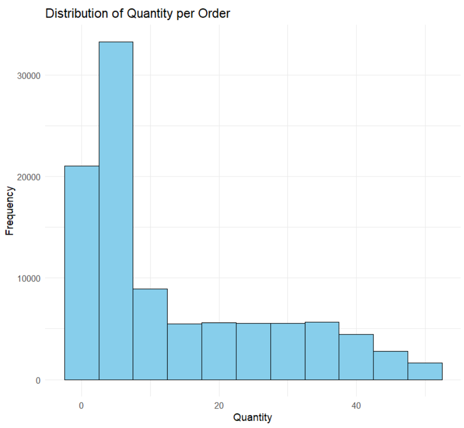


Figure 14: Distribution of Quantity per Order

For the final part of this analysis, the data files were combined to form a full database. This was done by linking the CustomerID in the sales file with the CustomerID in the customer database and linking the ProductID in the sales file with the ProductID in the combined products database.

The correlation matrix was constructed to evaluate the relationship between the numeric variables and identify points of interest. The correlation matrix is illustrated in the heatmap below. There is a strong positive correlation between SellingPrice and pickingHours due to the correlation coefficient 0,8412391125. This means that the higher the selling price of a product, the more time it takes to prepare the product. There is a moderate positive correlation, 0.5831669256, between pickingHours and deliveryHours. This means that products that take longer to prepare also take longer to deliver. These two relationships suggest that expensive products require additional processes to compensate for handling complexity.

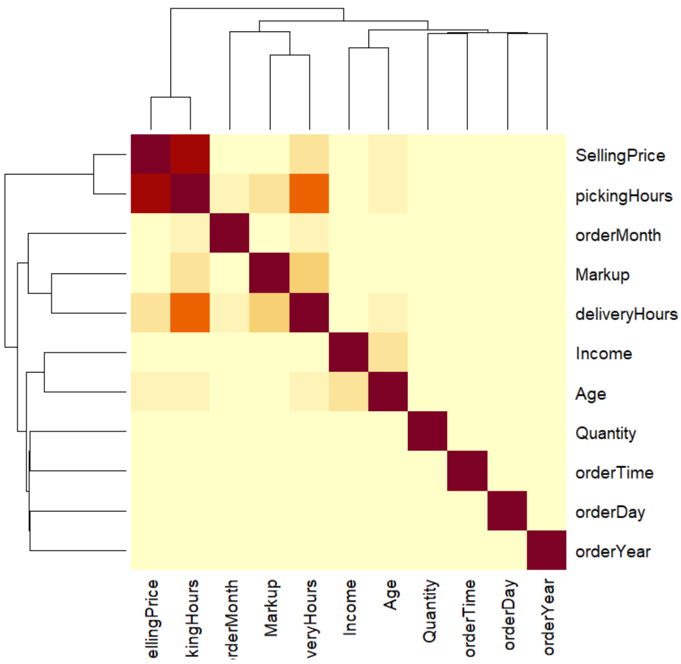
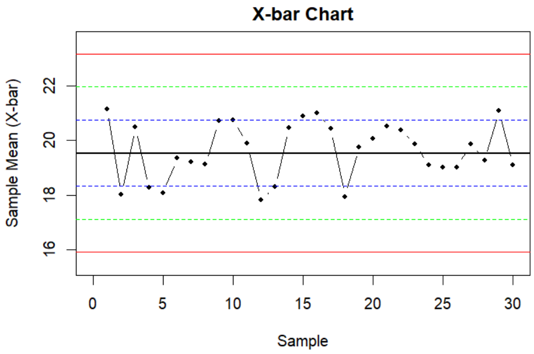
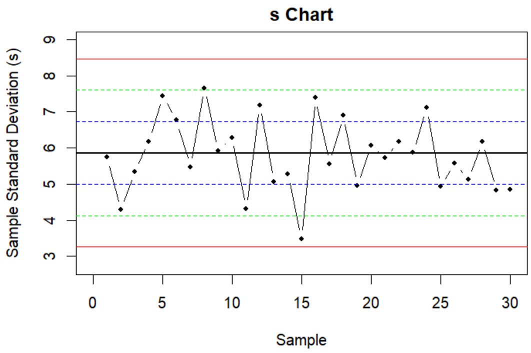


Figure 15: Correlation Matrix

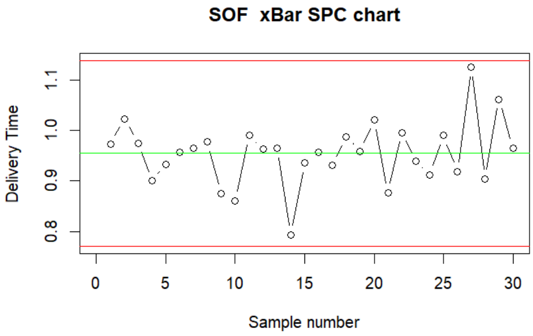
Part 3: Statistical Process Control (SPC)

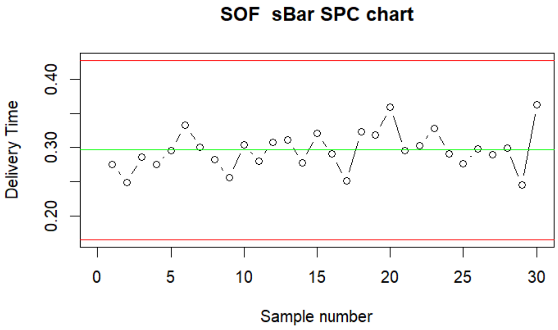
3.1 Control Charts for the first 30 samples



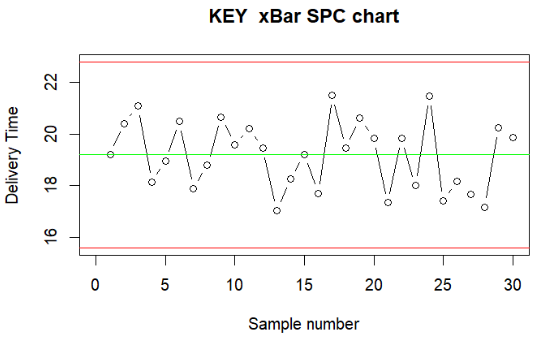


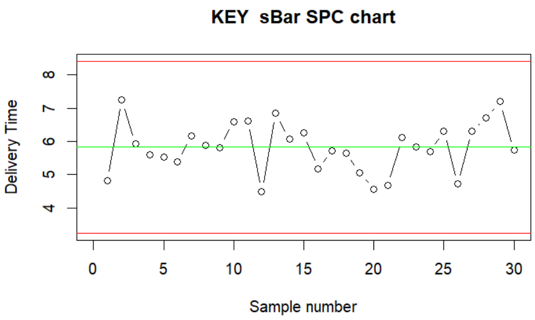
Product Type 1: Software



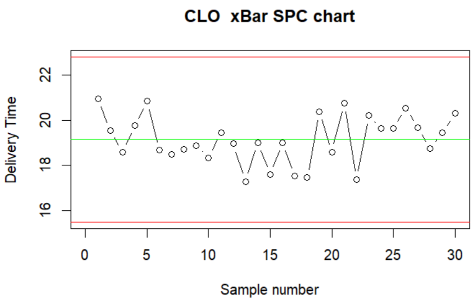


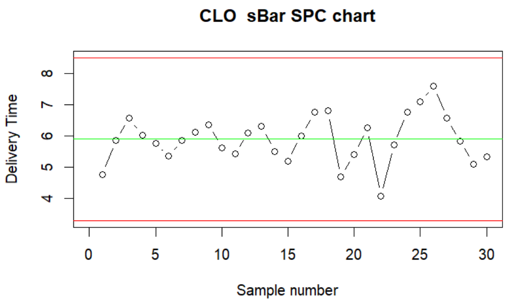
Product Type 2: Keyboard



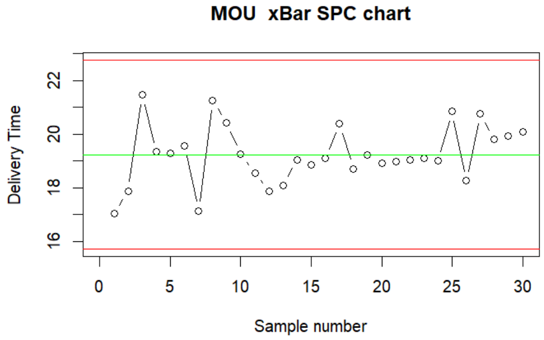


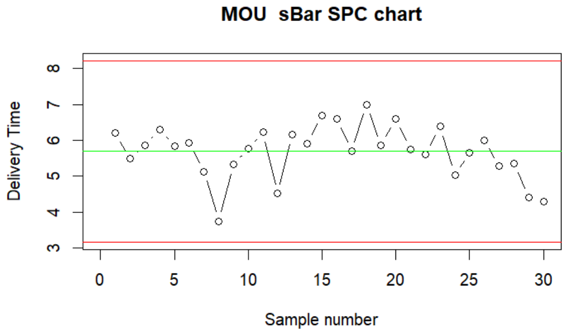
Product Type 3: Cloud Subscription



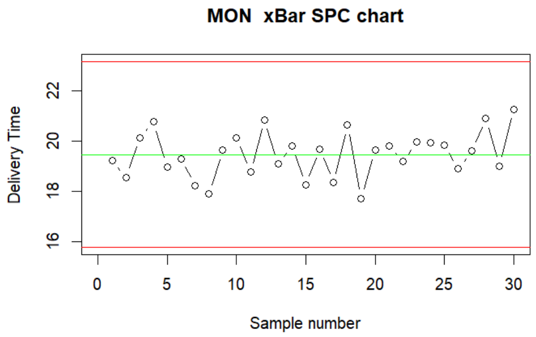


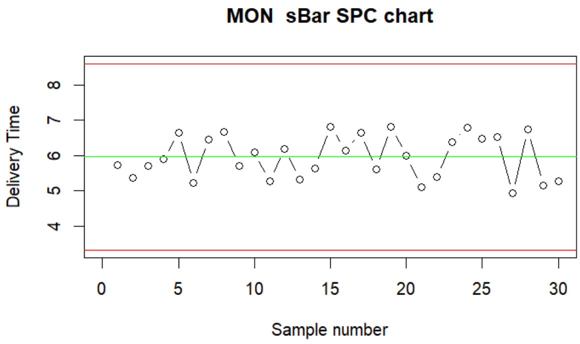
Product Type 4: Mouse



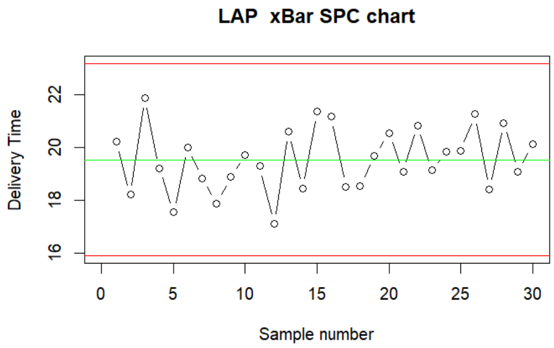


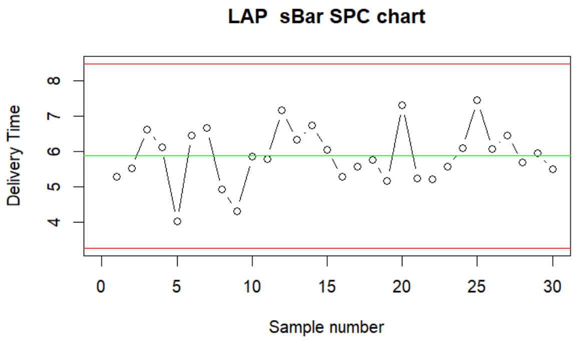
Product Type 5: Monitor





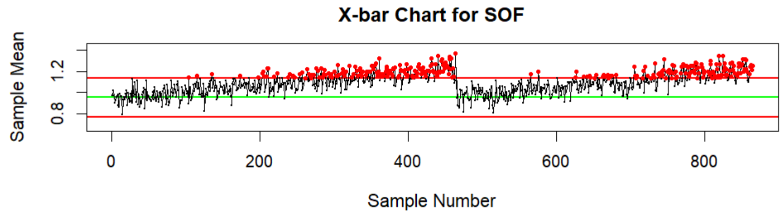
Product Type 6: Laptop

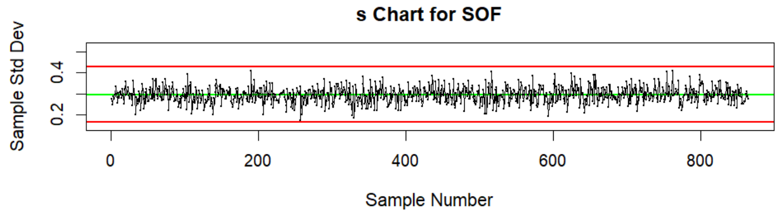




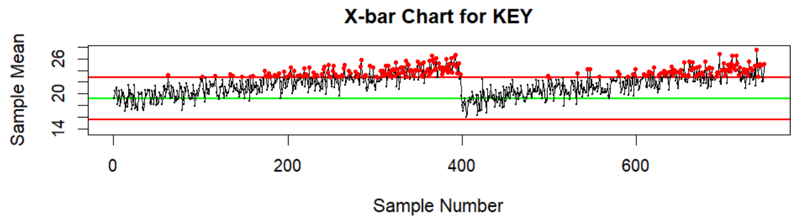
3.2 Control Charts for the rest of the data

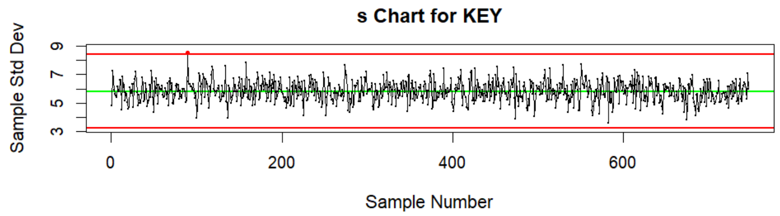
Product Type 1: Software



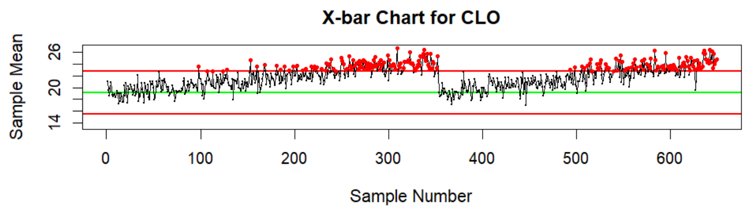


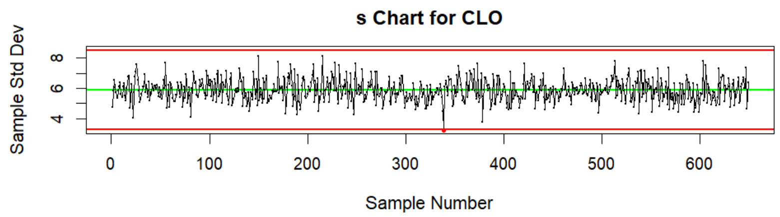
Product Type 2: Keyboard



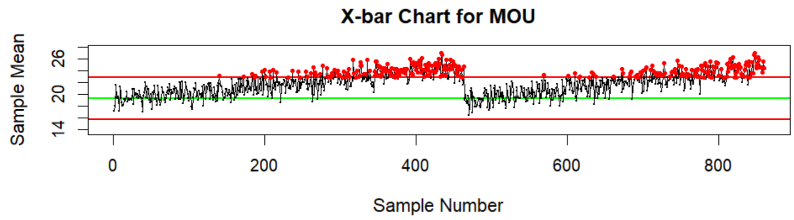


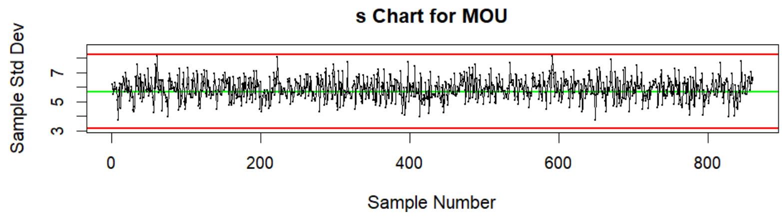
Product Type 3: Cloud Subscription



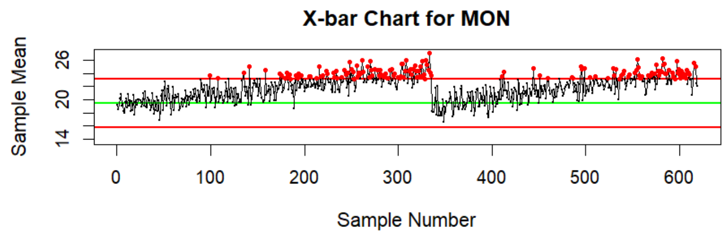


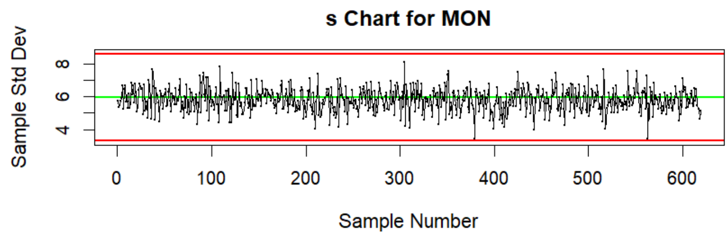
Product Type 4: Mouse



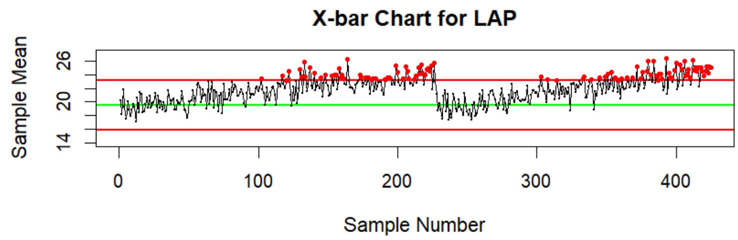


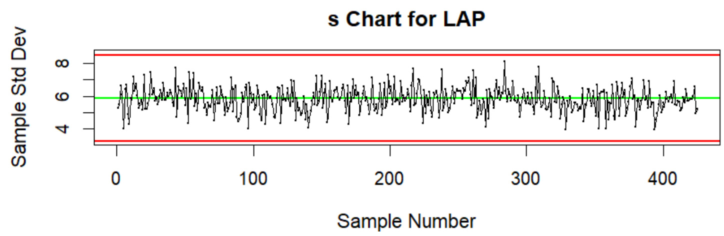
Product Type 5: Monitor





Product Type 6: Laptop





3.3 Process Capability indices

The analysis of process capability indices was conducted to determine if the delivery processes for the different product types are capable of meeting the Voice of the Customer (VOC), defined by a Lower Specification Limit (LSL) of 0 hours and an Upper Specification Limit (USL) of 32 hours. The following discussion is supported by the information in Table 2 below (Stellenbosch University, 2025a).

Software has a process capability index greater than 1 and less than 1.33, with Cpk = 1.0866, meaning that it is marginally capable of meeting the VOC. The Cpk is determined by the Cpl which means the process mean is very close to the LSL of 0 hours. This reflects consistently fast delivery times, which is supported by Figure 23 in Part 6. Software has a high process potential index, Cp = 18.1547, which demonstrates the extremely low process variation of the software delivery times.

The rest of the product types are classified as not capable of meeting the VOC because their Cpk values are all less than 1. Their Cpk values are determined by the Cpu values which indicates that the process average tends towards the USL instead of the LSL. The natural six sigma spread of the process is wider than the specification limit width, so the processes are not capable of meeting the 32-hour delivery target for all customers. Ultimately, they are not capable because their process variations are too high.

Table 2: Process Capability Indices

Product Type	Cp	Cpu	Cpl	Cpk	Capable
SOF	18.1547	35.2227	1.0866	1.0866	Marginally
KEY	0.9169	0.7298	1.1040	0.7298	Not
CLO	0.8972	0.7169	1.0774	0.7169	Not
MOU	0.9152	0.7254	1.1050	0.7254	Not
MON	0.8897	0.6999	1.0795	0.6999	Not
LAP	0.8988	0.6966	1.1009	0.6966	Not

3.4 Process Control Issues

This section identifies samples that show significant process control signals, indicating either out-of-control or stable in-control behaviour (Stellenbosch University, 2025a).

<u>Rule A</u>: s-chart points above UPPER control limit (UCL at +3σ)

This rule identifies one or more points that fall outside the upper control limit of the standard deviation chart.

- Findings:
  - KEY: 1 violation (Sample 89).
  - SOF, CLO, MOU, MON, LAP: No violations.

<u>Rule B:</u> The most consecutive samples of s between the ±1σ limits

This rule indicates good process control as there is stable variation.

Table 3: Rule B

Product Type	Longest Run	Sample Range
SOF	16	221 - 236
KEY	20	351 - 370
CLO	27	470 - 496
MOU	15	224 - 238
MON	21	46 – 66
LAP	17	239 - 255

<u>Rule C</u>: Four or more consecutive points on the Mean X-bar chart above the 2σ limit

This indicates a significant process shift or a sustained trend, it is a clear out-of-control signal

Table 4: Rule C

Product Type	Total Runs Identified	First 3 Run Start Samples	Last 3 Run Start Samples
SOF	286	208, 219, 220	859, 860, 861
KEY	222	178, 185, 186	741, 742, 743
CLO	227	180, 192, 193	644, 645, 646
MOU	229	249, 250, 251	855, 856, 857
MON	148	179, 203, 208	610, 615, 616
LAP	99	114, 115, 116	420, 421, 422

3.5 SPC Discussion

The analysis of the s-charts, which describe process variation, suggests that the consistency of the delivery process is stable. As noted in the Rule A findings, only one sample for the Keyboard product type breached the upper control limit on the s-chart. Furthermore, the Rule B analysis identified long, stable runs of in-control behaviour (within ±1 sigma) for all six product types. This indicates that the underlying process variation is predictable and stable.

The primary issue is not with process consistency but with the process average. The X-bar charts for all product types show significant and sustained out-of-control signals. This is highlighted by the Rule C findings, which identified hundreds of instances where four or more consecutive sample means fell above the +2 sigma limit. As established in the following section, the probability of an out-of-control signal being a false alarm is practically zero, confirming a true upward shift in the average delivery time from 2022 to 2023.

This upward shift in the process mean is the direct cause of the poor process capability findings. The processes for Keyboard, Cloud Subscription, Mouse, Monitor, and Laptop are all classified as "Not Capable" because their shifted process averages are too close to the Upper Specification Limit of 32 hours. The process is failing to meet the Voice of the Customer because it is no longer centred on its original, stable baseline.

In conclusion, the delivery process is not stable, and management intervention is required to investigate the root cause of this increase in average delivery times to bring the process back into a state of control and make it capable of meeting customer expectations. These investigations should focus on identifying the changes in the process between 2022 and 2023. This can be done by interviewing staff about any changes in procedures, reviewing staffing levels in the warehouse to determine if they have optimal number of workers on shift at a time and investigating if their logistical partners have changed or if their performance has significantly declined.

3.6 Taguchi Loss

The SPC conclusions and recommendations align with the core message of Taguchi Loss. They both agree that process variation is the enemy of quality. The SPC analysis found that the delivery process is fundamentally unstable and out of statistical control. The essence of Taguchi is that quality is achieved by reducing variation to create a stable, predictable process. Therefore, the recommendation to investigate and bring the process back into a state of control perfectly aligns with this principle. Both frameworks ultimately agree that this unstable, shifting process is the true root cause of the loss and must be corrected

Part 4: Risk, Data correction and optimising for maximum profit

4.1 Likelihood of making a Type I (Manufacturer's) Error

A Type I error occurs when H_0 is true, but our rule gives a signal, causing us to reject H_0 by mistake.

H_0: The process is in control, stable and centred at the centreline calculated from the first 30 samples.

H_a: The process is not in control or stable. The mean has shifted, or variation increased.

<u>Rule A</u>: The probability of being beyond 3σ

Type I Error = P(sample > UCL | process is in control)

$\alpha = P(|Z| > 3) = [1 - \text{pnorm}(3)] \approx 0.00135$

There is a 0.135% chance that a single sample will fall above the +3σ UCL just by random chance, even when the process is in control.

<u>Rule B</u>: Probability of a sample being between the ±1σ control limits

Computed by the code: `pwithin1sigma <- pnorm(1, 0, 1) - pnorm(-1, 0, 1)`

The probability is 0.6827.

Therefore, we expect 68,27% of all in control samples to fall within the ±1σ control limits

<u>Rule C</u>: The probability of 4 consecutive samples beyond +2σ

P(one sample > +2σ) = 1 - pnorm(2) ≈ 0.0228

Type I Error = (0.0228)^4 ≈ 0.00000027

The probability of getting this signal as a false alarm is 0.000027%. This makes it a very strong and reliable signal that the process mean has shifted and is no longer in control.

4.2 Likelihood of making type II (Consumer's) Error

Beta is calculated by calculating the probability that the samples means still fall within the original control limits as this will result in failing to detect the shift. This was done using the following line of code.

```
pnorm(UCL, mean = mu_actual, sd = sigma_xbar_new) - pnorm(LCL, mean = mu_actual, sd = sigma_xbar_new)
```

Beta = 0.8411783

Therefore, there is a 84,12% probability of a type II error.

4.3 Products Data Correction and Re-Analysis

The descriptive analysis in Part 1 is misleading due to the data quality issues. After making the required corrections to the data, we now know our analysis is based on reliable and consistent data. The data now represents the true state of the company's product pricing.

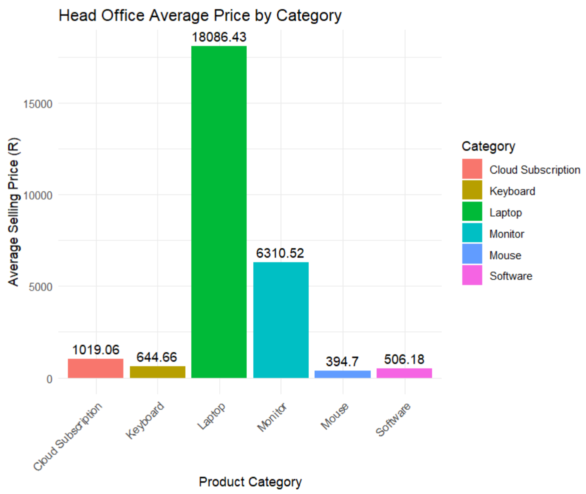


Figure 16: Head Office Average Price by Category

The bar chart in Figure 16 gives an accurate depiction of the average selling price for each product type. The Laptop category is a significant outlier with an average selling price of R18086,43. The other product categories are all relatively low-cost items with Mouse being the least expensive product category with an average selling price of R394,70.

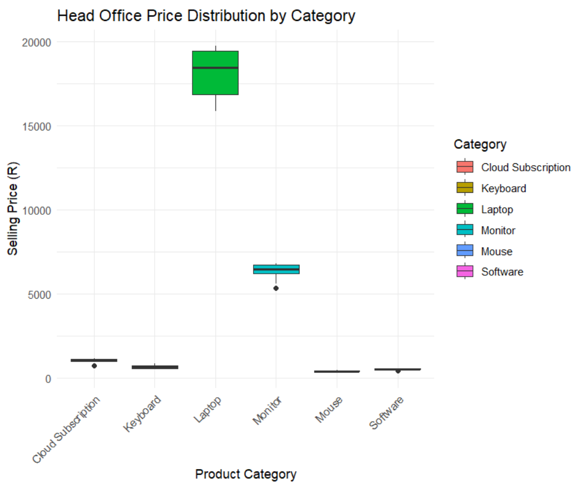


Figure 17: Head Office Price Distribution by Category

The boxplot in Figure 17 above, provides insight into the distribution and variability of the average selling price in each product category. The Laptop category has the widest interquartile range and overall spread, demonstrating extreme price variability. The Monitor product category shows that the average selling price in that category is also highly variable. The rest of the product categories have consistent pricing with almost no outliers. The primary finding from the re-analysis is the clear price segmentation of the company's products, with Laptops and Monitors representing high-value, high-variability items, and the other four categories representing low-value, low-variability items.

Part 5: Optimising Profit

5.1 Shop 1 Analysis

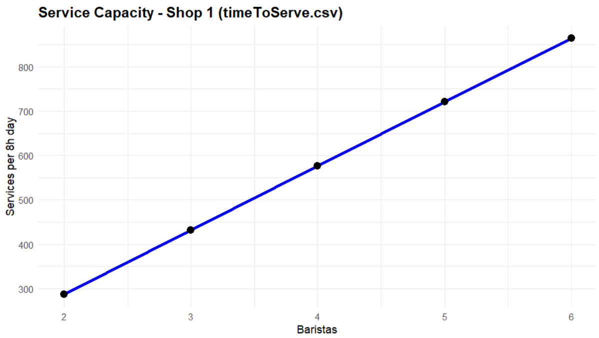


Figure 18: Shop 1 Service Capacity

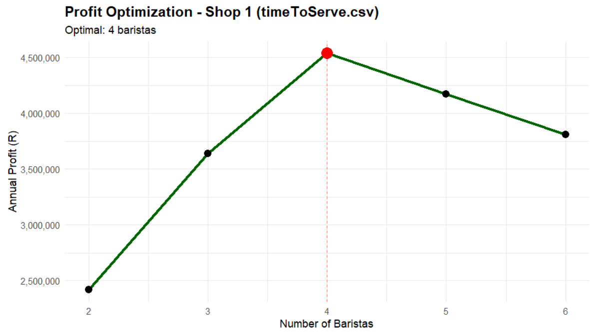


Figure 19: Shop 1 Profit Optimization

This optimisation problem balances the cost of personell while maximising the revenue generated annually by Shop 1. The optimal number of baristas is 4, shown explicitly in the profit optimization graph, as 4 baristas is the global maximum and the point of diminishing return. This model also assumes 200 000 transactions annually. This equates to an average daily demand of 548 customers/day. On the Shop 1 Service Capacity graph, service capacity increases as baristas are added. Therefore, the shop is demand-constrained at 4 baristas because it slightly exceeds the average daily demand, resulting in a 100% service reliability. Shop 1 expects R 4 540 000 annual profit.

5.2 Shop 2 Analysis

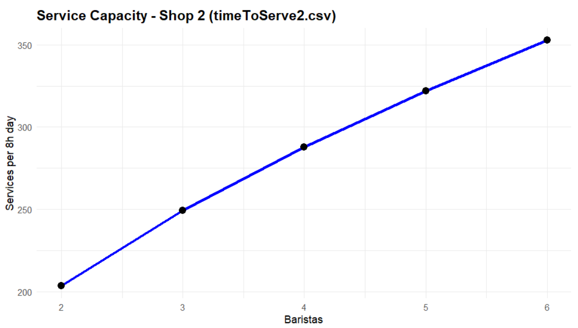


Figure 20: Shop 2 Service Capacity

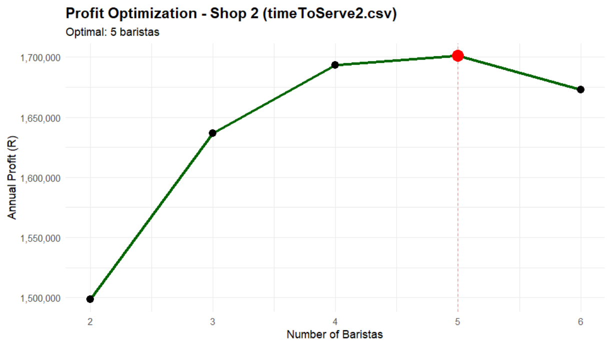


Figure 21: Shop 2 Profit Optimization

The profit optimisation graph shows that the annual profit will increase when the number of baristas is increased, peaking at 5 baristas as the point of diminishing return. As can be seen in the Shop 2 Service Capacity graph, shop 2 has a more constrained capacity. The maximum service capacity of shop 2 is 350 customers/day. At the optimal level of 5 baristas, the shop can only serve 322 customers per day, which is far less than the 548 service capacity of shop 1. This results in the expected service reliability being 58.77%. Five baristas create the highest possible annual profit of R1 701 098. This is still the optimal solution even though 41% of customers are not served, this is the most profitable solution for the business.

Part 6: DOE and MANOVA or ANOVA

The SPC analysis in Part 3 showed that the delivery time processes are out of control for most of the product types. The X-bar charts indicated numerous samples above the UCL, and the Rule C analysis confirmed a sustained upward trend in delivery times. This visible process shift appears to correspond to the change from 2022 to 2023 in the dataset. Therefore, this section will use an Analysis of Variance (ANOVA), to statistically test the hypothesis that there is a significant difference in mean delivery times between 2022 and 2023 (sthda, 2022). The two-way ANOVA is used to test the effect of orderYear and ProductType on deliveryHours. The following line of code in Figure 22 executes the two-way ANOVA.

```
anova_model <- aov(deliveryHours ~ orderYear * ProductType, data = future_data)
summary(anova_model)
```

Figure 22: Two-way ANOVA in rstudio

Table 5: ANOVA Summary

	Df	Sum Sq	Mean Sq	F value	p-value
orderYear	1	139	139	4.707	0.0300
ProductType	5	7022849	1404570	47696.755	<2e-16
orderYear:ProductType	5	273	55	1.852	0.0991
Residuals	99988	2944438	29		

From the summary table above, we can make the following conclusions. The orderYear has a p-value of 0.03 which is very small (<0.05) and proves that there is a statistically significant difference in mean delivery times between 2023 and 2022. The ProductType has an extremely small p-value, which proves that different product types have statistically different mean delivery times. The interaction between orderYear and ProductType has a large p-value which indicates that the interaction is not statistically significant. Therefore, the change in delivery time from 2026 to 2027 is statistically consistent across all six product types. This means that although the SPC charts may visually depict one product having a larger shift than another, the difference in the size of the shift is not statistically meaningful. Therefore, the change in delivery time from year to year affected all six product types in a similar way.

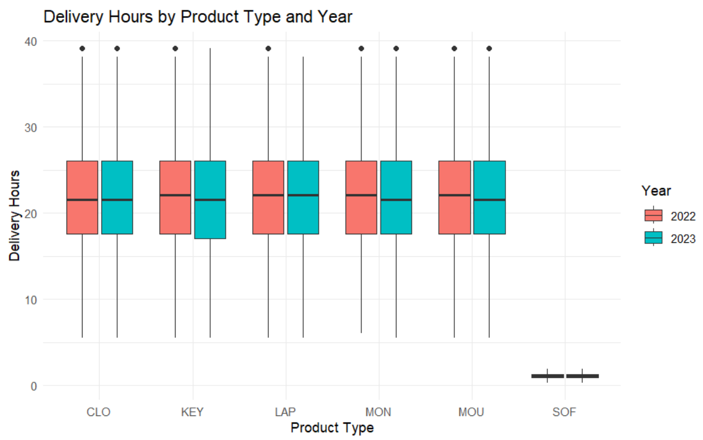


Figure 23: Delivery Hours by Product Type and Year

The figure above confirms the findings of the two-way ANOVA test. This figure visually depicts the significance of the ProductType variable as the SOF category's delivery time is consistently near zero while the other delivery times of the other categories show a higher distribution. For each ProductType, the distributions are almost identical which confirms that no single ProductType behaved differently from the others between the two years. The ANOVA test indicates there is a statistically significant difference between the years, however in context, the difference in delivery time is extremely small.

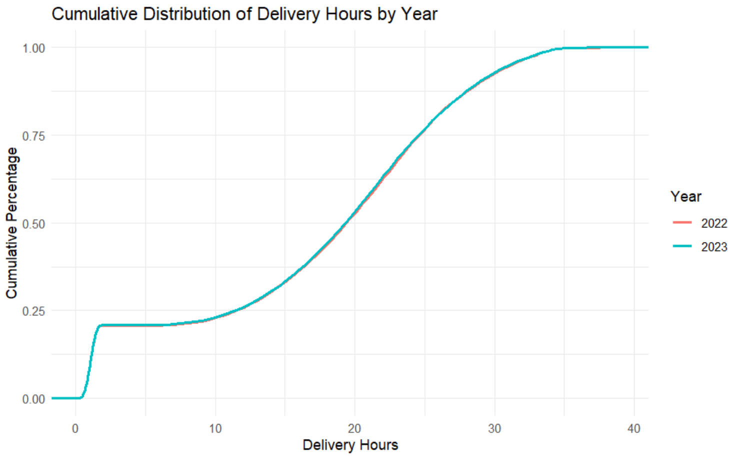


Figure 24: Cumulative Distribution of Delivery Hours by Year

The cumulative distribution plot of delivery hours by year shows the lines of the two years almost perfectly overlapping, indicating that the delivery performance was practically identical for both years. It can also be seen that approximately 20% of deliveries were completed within the first few hours in both years. While the orderYear p-value indicated that there is a statistically significant difference in the mean delivery times between the two years, this plot does not visually support this claim. This is likely due to the ANOVA test being sensitive enough to detect a minor difference due to the large sample size. The practical performance between the two years is minor.

Part 7: Reliability of service

7.1 Expected days of reliable service per year

Total reliable days: 96 + 270 = 366 days

7.2 Optimising Profit

As can be seen in Figure 25 below, 17 is the optimal number of personnel to assign as it minimises the cost to the car rental agency.

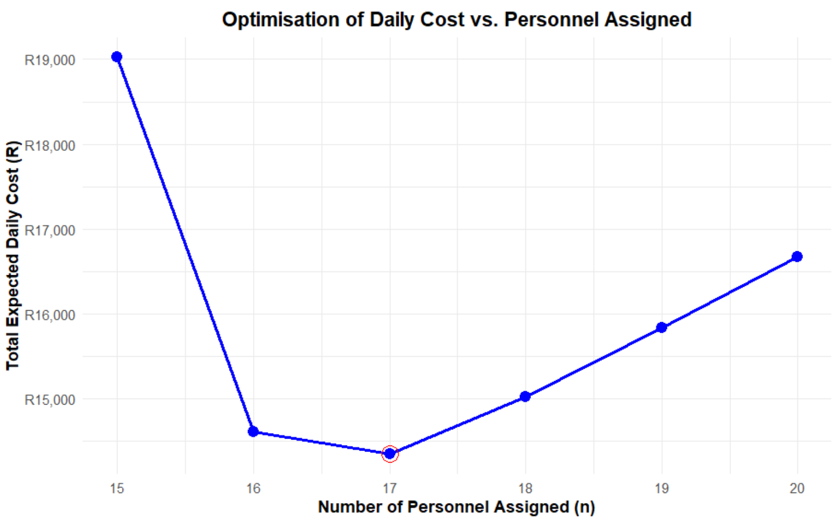


Figure 25: Car Rental Agency Optimisation

Conclusion and Recommendations

The critical data integrity issue between the local branch and head office product databases is the most urgent issue management must address. These outdated and inconsistent databases can cause future problems for the business when synthesizing data for analysis or reporting sales and other critical information to shareholders. This can also lead to inefficiencies like stockouts as the head office is unaware of the stock level of the unknown items at the local branch and therefore is unable to replenish them. The local branch is also potentially missing sales opportunities by not carrying the full product range from the head office. My recommendation to upper management is to remedy this situation to ensure all databases are updated and relevant. Therefore, the initial analysis in Part 1 provides an incorrect and misleading view of the company's profits. Upon data correction detailed in Part 4, it is found that Laptops and Monitors are the high-value, high-variability categories, not Software, as was previously assumed.

Furthermore, the time-based analysis conducted in Part 2 illustrates the declining total sales from 2022 to 2023. Daily and monthly sales patterns remained consistent meaning that the decline is due to a decline in performance rather than a seasonal anomaly. Management must investigate the reason for this by engaging with customers.

Another major concern is the sustained upward shift in average delivery time, detailed in the SPC analysis in Part 3. The delivery process is out-of-control and not capable of meeting the 32-hour Upper Specification Limit for all product types except software, which is characteristically independent of the typical logistical factors of a delivery process. Management should launch a formal investigation, such as a Six Sigma/DMAIC project to identify and eliminate the root cause of the upward shift in average delivery time. This formal investigation should prioritize high value items such as Laptops and Monitors. The ANOVA in Part 6 statistically confirms that this shift is linked to the orderYear, and Part 2 established a clear correlation between a product's SellingPrice, its pickingHours, and its deliveryHours.

Finally, the optimization models in Parts 5 and 7 provide clear, data-driven solutions for staffing. These models demonstrate that optimal staffing is not one-size-fits-all, as seen by the different strategies for the demand-constrained Shop 1 versus the capacity-constrained Shop 2, and the cost-minimization model for the car rental agency.

In conclusion, the business should use this report to re-align their business strategy using corrected data and updated analysis. The business strategy should focus on high-value categories such as Laptops and Monitors and not low-value, low-variability items in the other categories to maximise profit, customer retention and customer satisfaction.

References

Engineering Counsel of South Africa, 2025. *Project ECSA 2025 Part 1, 2, 3, 4, 5, 6, 7*. [pdf] Stellenbosch University. Available at: ProjectECSA2025Part1234567.pdf [Accessed October 2025].

sthda, 2022. *MANOVA Test in R: Multivariate Analysis of Variance - Easy Guides - Wiki - STHDA.* [Online] Available at: http://www.sthda.com/english/wiki/manova-test-in-r-multivariate-analysis-of-variance

Stellenbosch University, 2025a. *QA344 Formula Page*. [pdf] Stellenbosch University. Available at: QA344FormulaPage.pdf [Accessed October 2025].

Stellenbosch University, 2025b. *SPC Basics*. [pdf] Stellenbosch University. Available at: SPCBasics.pdf [Accessed October 2025].