Quality Assurance

ECSA Project

Name: Emily Roberts

Date: 24/10/2025

Student Number:

26924579

Table of Contents

Introduction

This report contains the results of a data analysis study that was conducted based datasets that were provided. These datasets included: customer_data, prod_headoffice_2025, products_data, Products_data2025, sales2022and2023, sales2026and2027, timeToServe, and timeToServe2. The data analysis aimed at discovering significant patterns within the datasets that could prove useful for the company's future operations.

The analyses that were undertaken included a customer demographic analysis, product analysis, statistical process control analysis, process capability analysis, risk, data correction, profit optimisation analysis, ANOVA, and service reliability study. Overall, the findings from these studies can assist the relevant companies in making necessary changes and improvements in the future.

Customer analysis

The customer data provided the customer ID, age, gender, city, and income for 5000 clients.

Customer gender

From the client gender distribution in Figure 1, it can be seen that there are only a few more females than males. However, this difference is insignificant. The “other” category may be an indication that the survey was not completed, the client preferred not to stipulate their gender, or an error during data collection. These instances could be removed. This is unnecessary since the client's gender will seemingly not contribute to any patterns in the data.

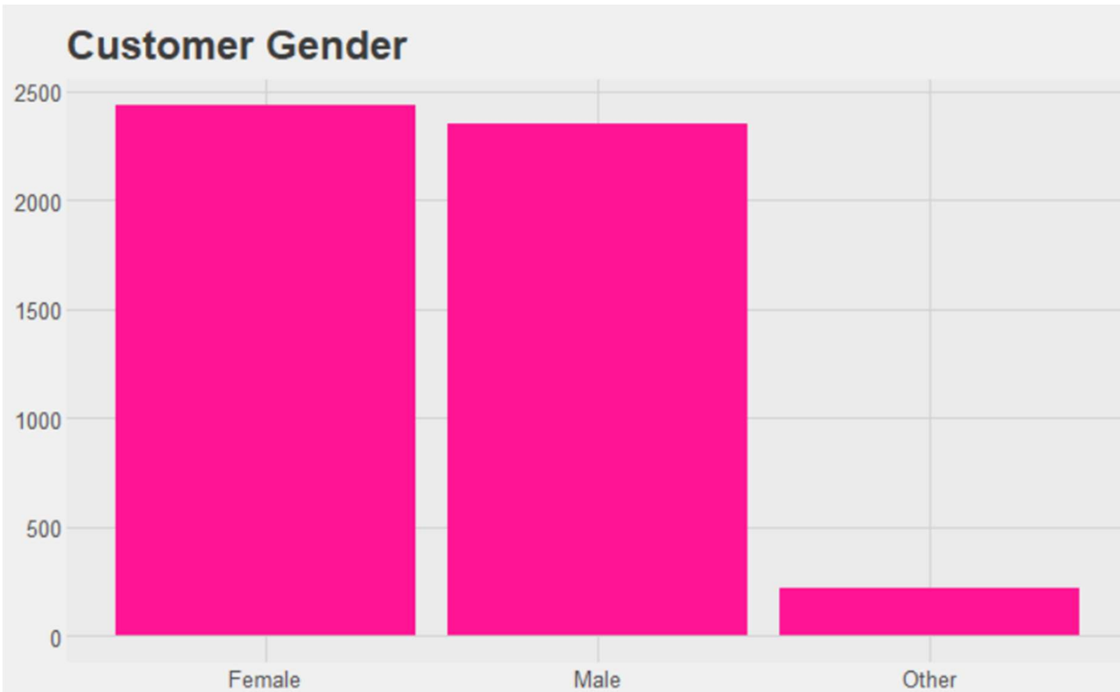

Figure 1: Customer Gender

Income distribution by gender

The female and male genders have a wide range of incomes (which are notably larger than the “other” category), but have a similar mean. The mean income for the female clients is only marginally larger than males. The 1.5 IQR range, first and third quartiles are similar for all categories. This means that the upper and lower tails are comparable. The main difference between the 3 categories is the mean values. The incomes are seemingly balanced, which can be seen by the significant distribution overlapping in Figure 2. It is clear that gender would not majorly influence income-related patterns. The deductions from the box plot are further supported by the density distribution graph seen in Figure 3.

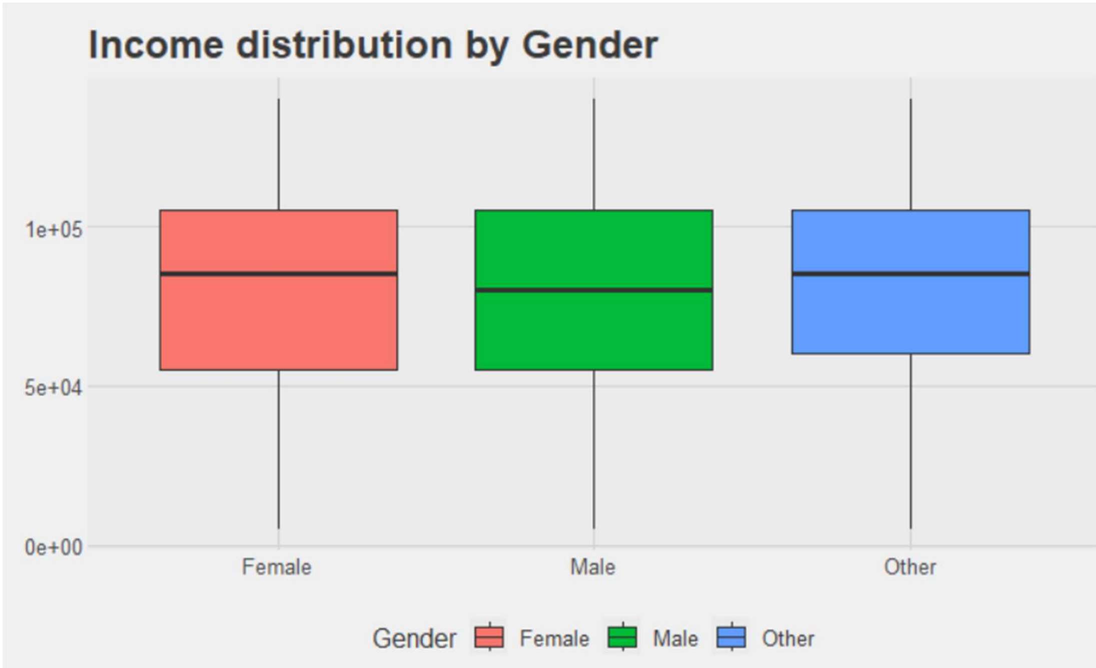

Figure 2: Boxplot for income distribution by gender

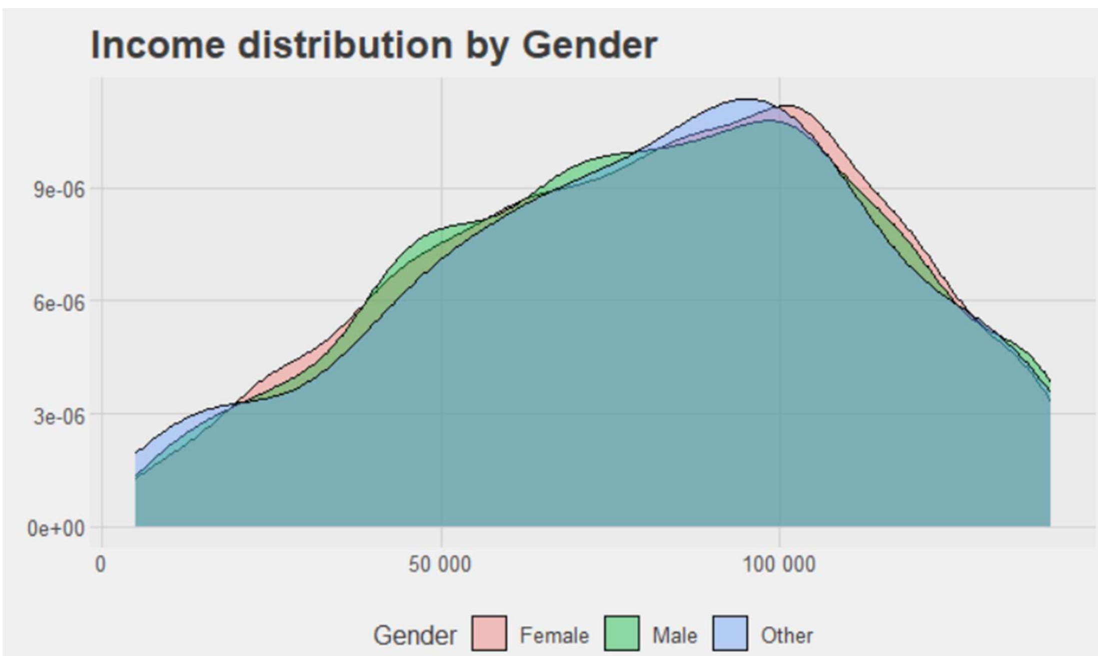

Figure 3: Density distribution graph for income distribution by gender

Customer age

By investigating Figure 4, it can be seen that the majority of the customers that are registered with the company are in the age group that ranges from 30-35 years. The number of customers gradually decreases as the age increases. The decrease is linear after the age of 75. The oldest client who is registered is over 100. This would suggest that the company is mainly aimed at a younger demographic. This information could be beneficial when the company is looking at marketing strategies.

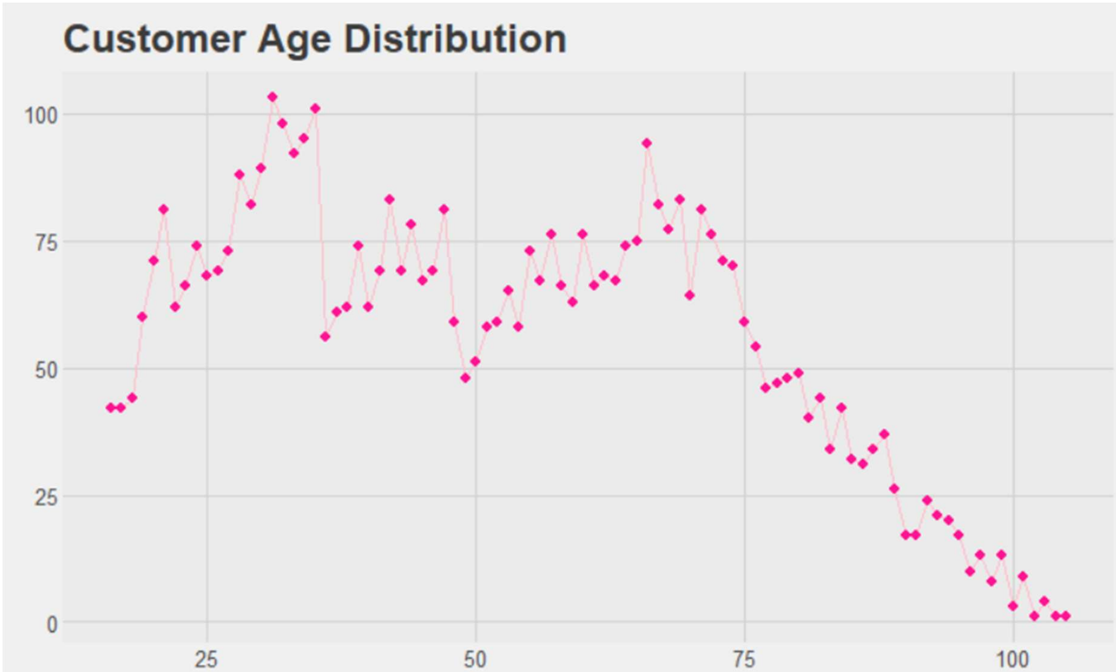

Figure 4: Customer age distribution

The incomes for the various age groups can be investigated in Figure 5. The average income increases dramatically for the age group in the range of 35-60. This is understandable, as this age group has gained work experience and would be the group that would be the most densely employed and encounter promotions. The younger age group would be paid less as they are new in the workplace, and do not get high salaries. After the age of 6o (which is close to the age of retirement), the salaries start to decrease again. Fewer people are working in this age group, therefore also lowering the average. The trends seen in the graph support the idea that the middle-aged group will have the highest purchasing power. Thus indicating that the age per income distribution would provide useful patterns for the company's future marketing strategies and planning.

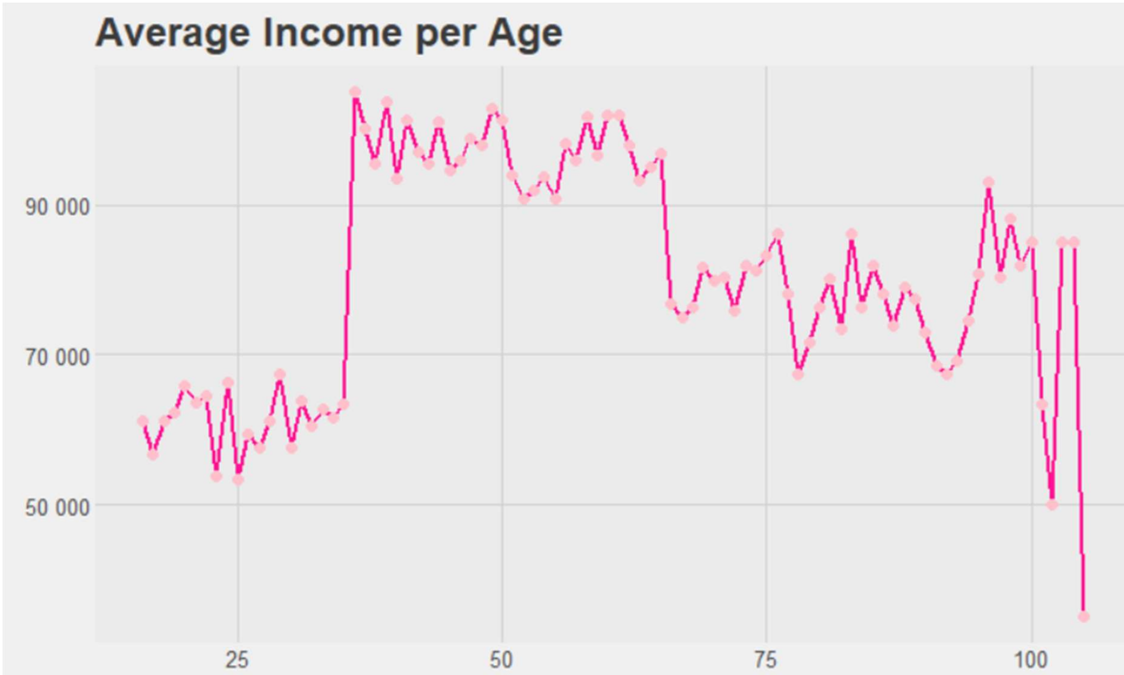

Figure 5: Average income for the ages

The average incomes for the cities are very similar, however Chicago and Miami is it slightly higher. Since the information does not contain any real patterns that would be of interest, it is somewhat irrelevant.

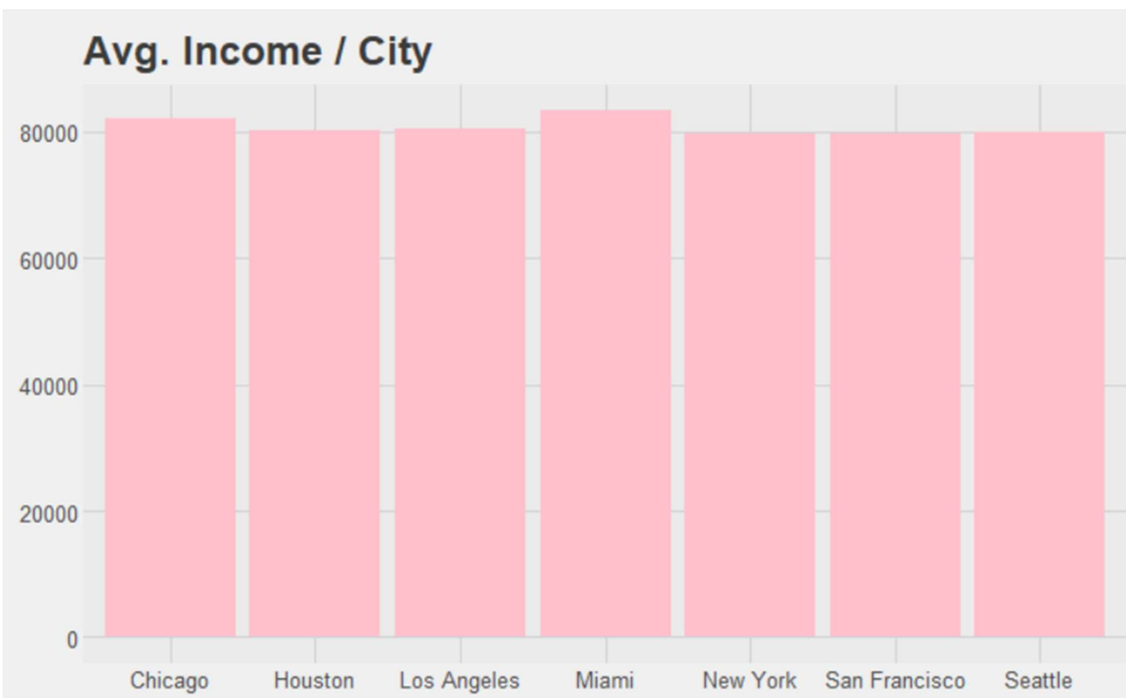

Figure 6: Average income per city

Conclusion

Overall, it can be seen that gender does not affect the trends analysed. This contrasts with the age category that provided insightful information with regard to the purchasing trends. It showed that purchasing power was strongly related to income. Middle-aged people who had the highest incomes were those who had the strongest purchasing power. This demographic should be intentionally targeted by the company to try and maximise profits.

Product analysis

The product analysis will include graphical interpretations of the information gathered in the products and the product's head office datasets.

Selling price per category

The selling prices for each product category can be seen in Figure 7. Cloud subscription has the lowest, while a laptop has the highest. The increasing prices of each category is logical as the value of the products in each category increases from the bottom up. This distribution does not show any new or revolutionary patterns that could be useful. The conclusions from this graph could have been logically achieved.

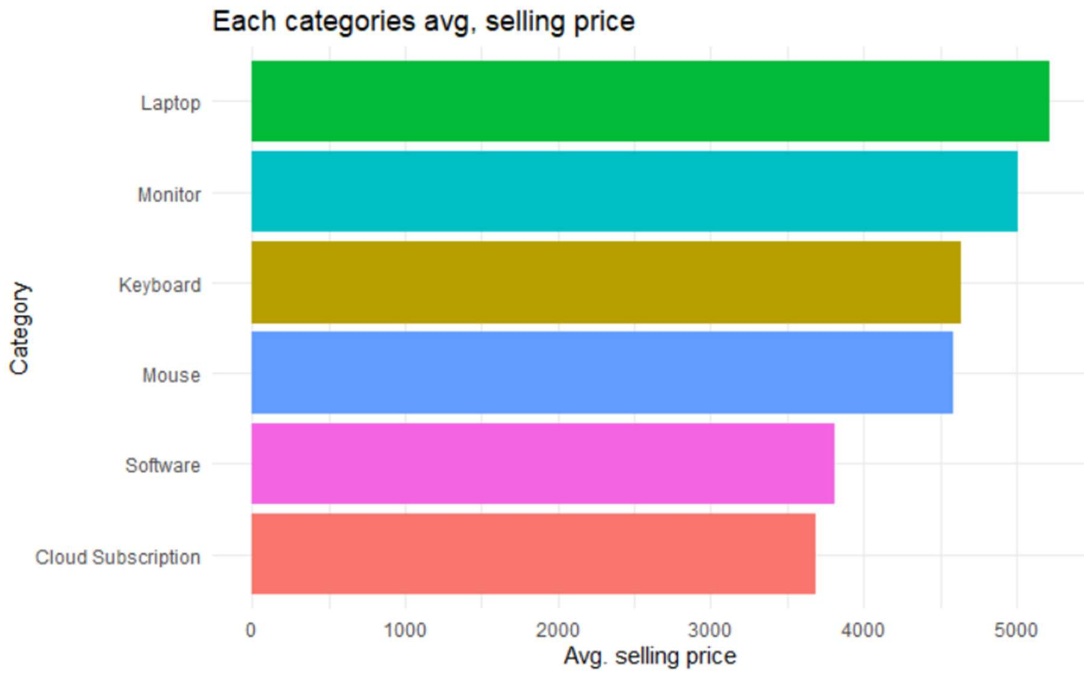

Figure 7: Average selling price per product category

In Figure 8, it is seen that there is much less variation in the head office selling prices. The laptop has the lowest, and the mouse has the highest. This shows that the head office has different pricing strategies that could be for various reasons. This could be because of promotional discounts, or internal cost allocations. This information could be of use.

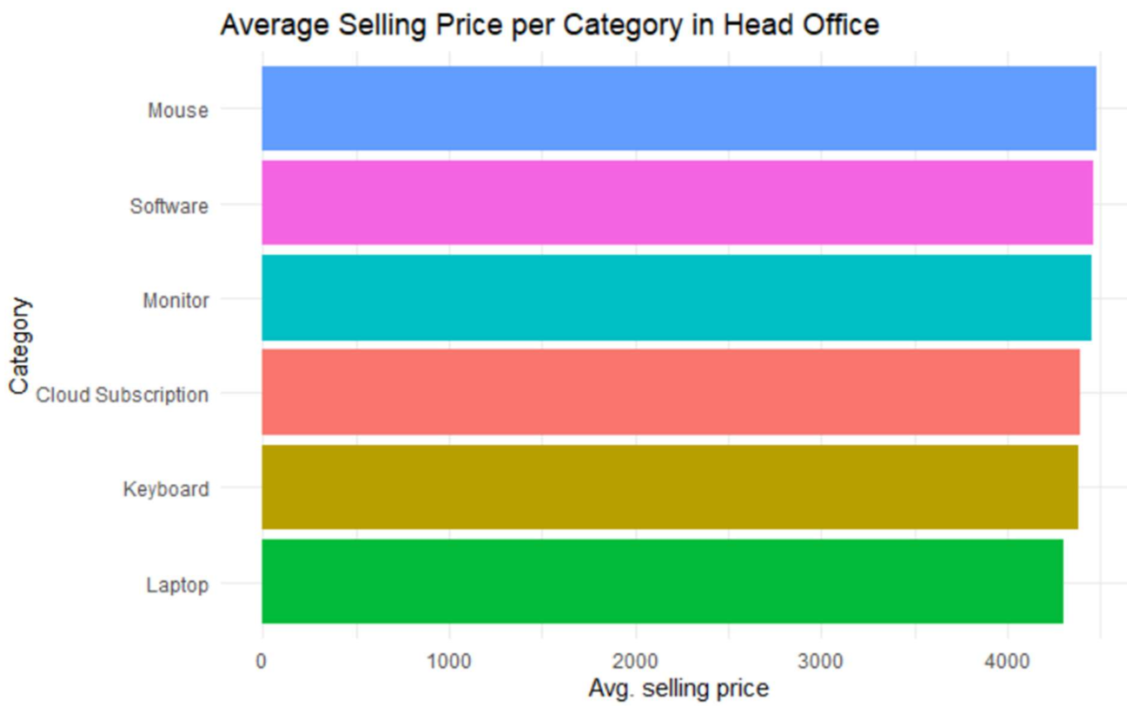

Figure 8: Average selling prices per category in head office

The same conclusions can be deduced from Figures 9 and 10. The upper quartiles of each category vary much more in Figure 9 than they do in Figure 10. Figure 10 shows a more standardised pricing approach than Figure 9.

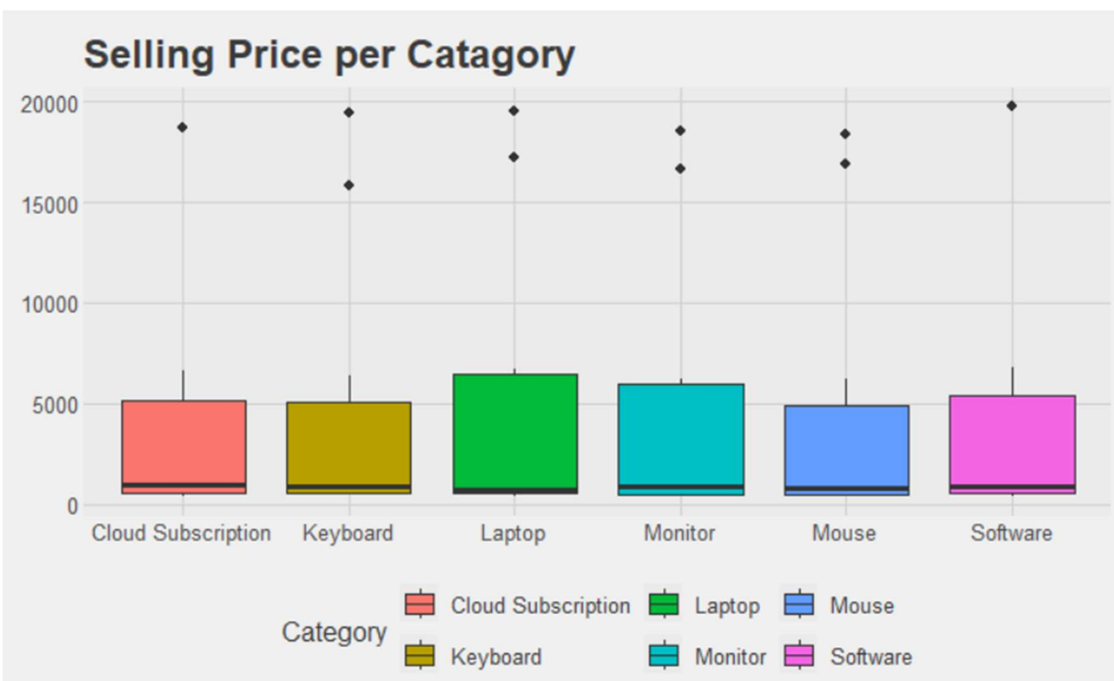

Figure 9: Selling price per product category

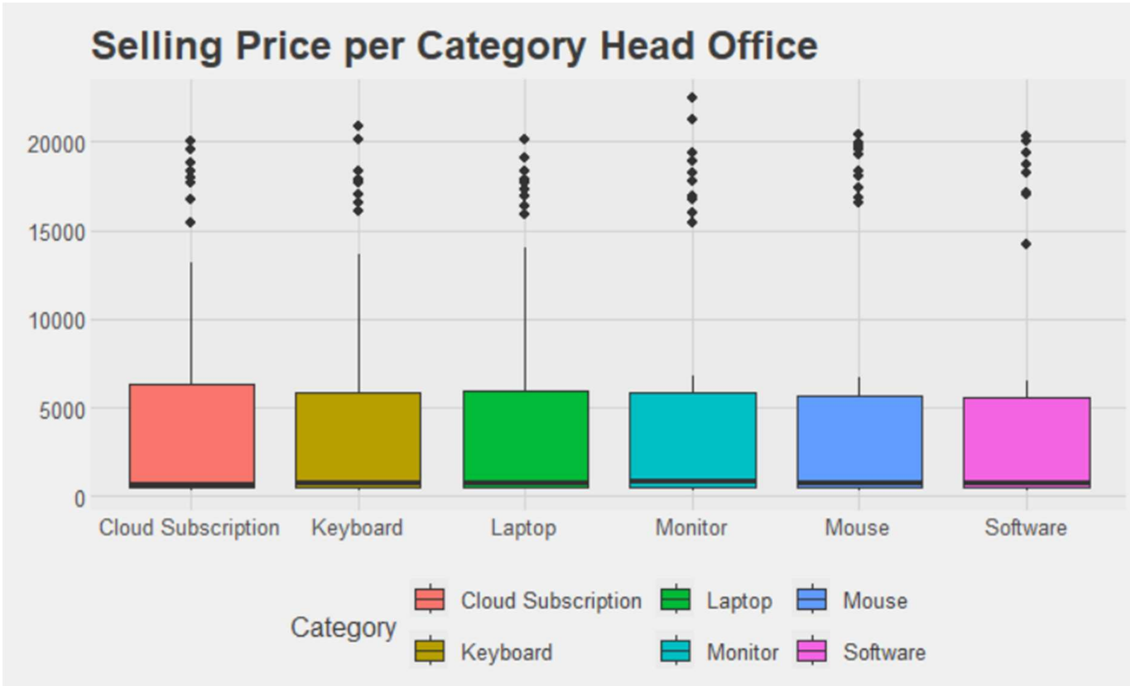

Figure 10: Selling price per product category in the head office

Distribution of markup

In Figure 11, it can be seen that there is a rather uniform distribution. This indicates that the products have an even and wide range of markups. There is no concentration on a single value. Therefore, pricing strategies are varied across the different categories. However, in Figure 12, there is a normal distribution. Therefore, the is a stronger concentration on an average markup value. This means that there are probably extreme markups, but most products are priced consistently, showing that there is a very standardised pricing approach.

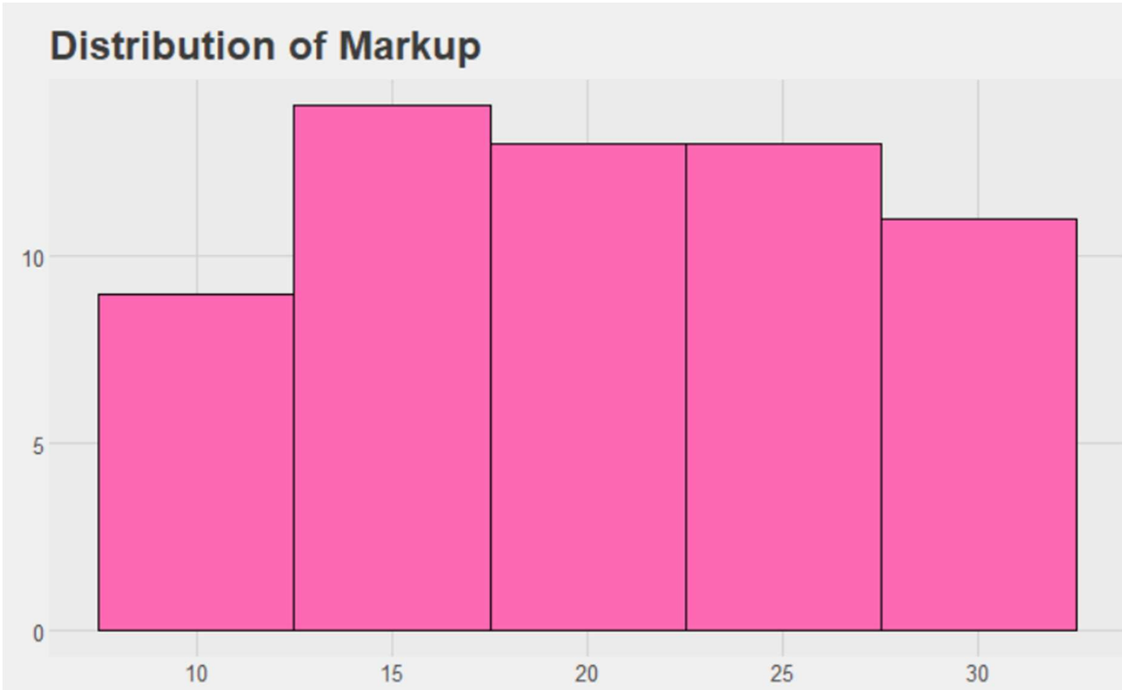

Figure 11: Markup distribution

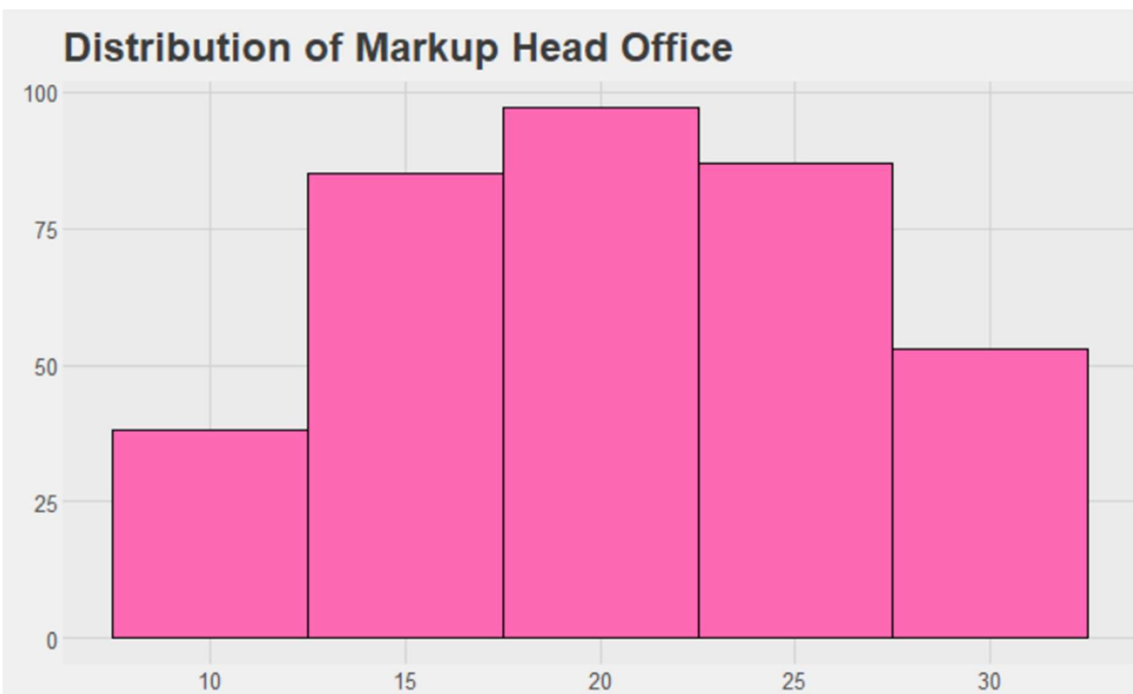

Figure 12: Markup distribution in head office

The boxplots reinforce the bar graphs. In Figure 13, there are wider interquartile ranges and more variation between categories. While the grouping is much tighter in Figure 14. Consistent pricing decisions in the office are highlighted in contrast to the market variability from the dataset.

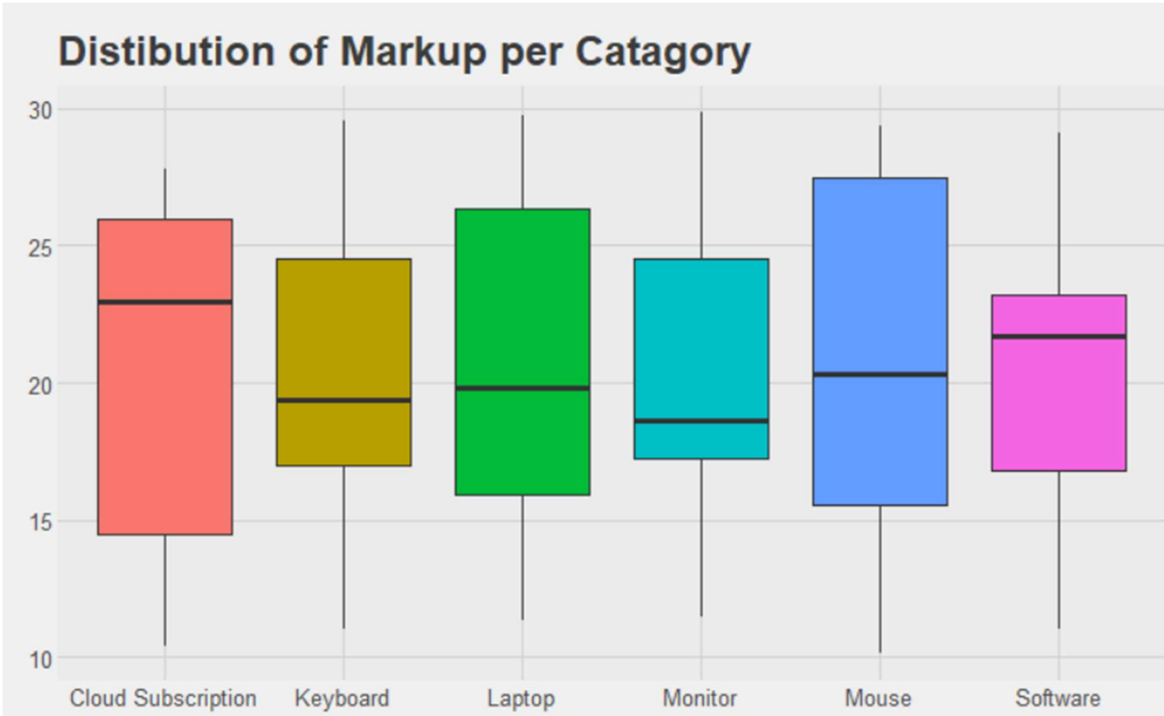

Figure 13: Markup distribution in boxplot

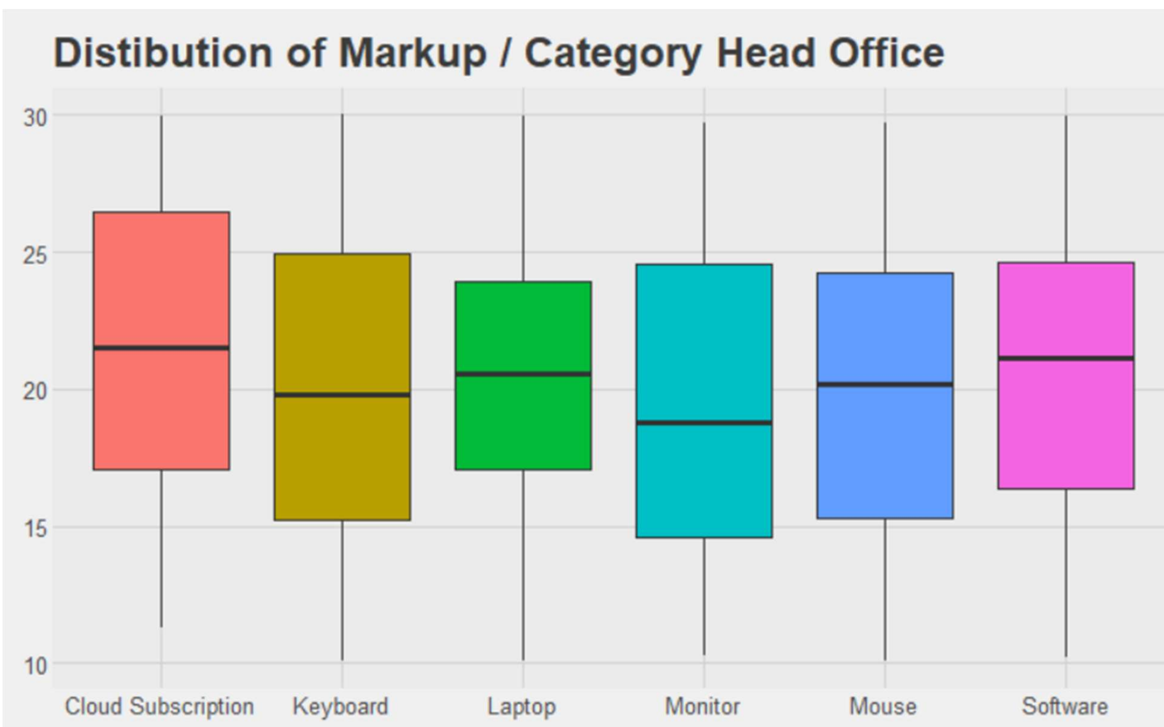

Figure 14: Markup distribution in boxplot (head office)

Profit

Selling price trends are mirrored by the distributions of profit. This is seen in Figure 15 as products such as laptops have much higher profit, and cloud subscriptions are much lower. However, as seen in the selling price section, there is a much more uniform spread across the categories in Figure 16.

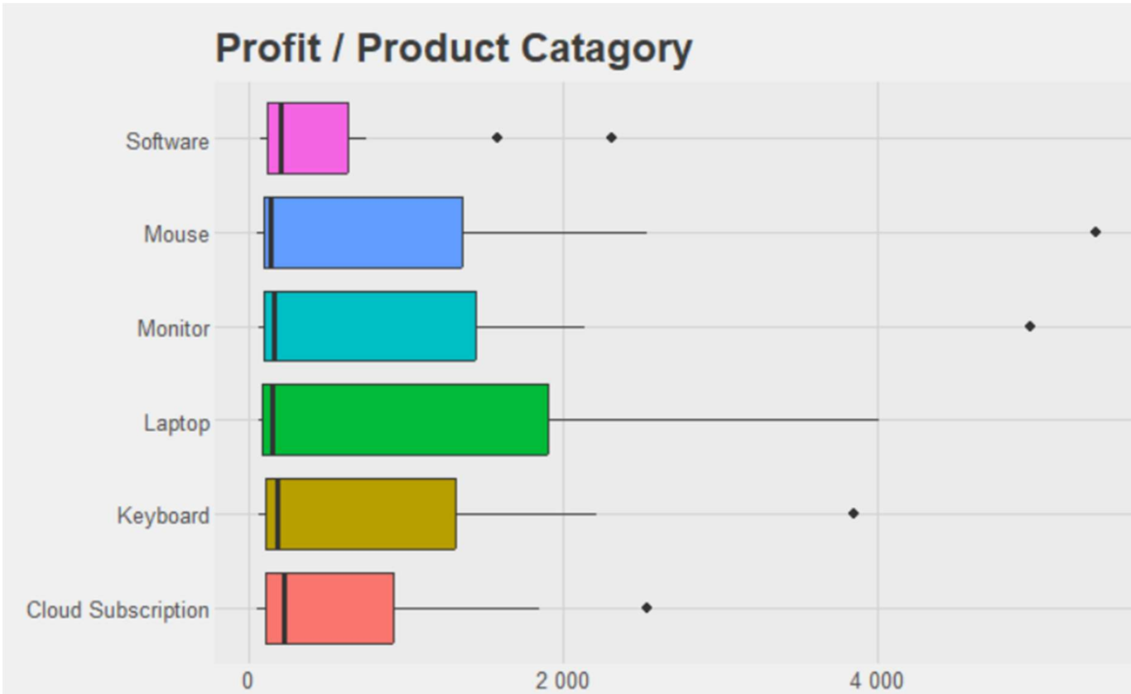

Figure 15: Profit per product category

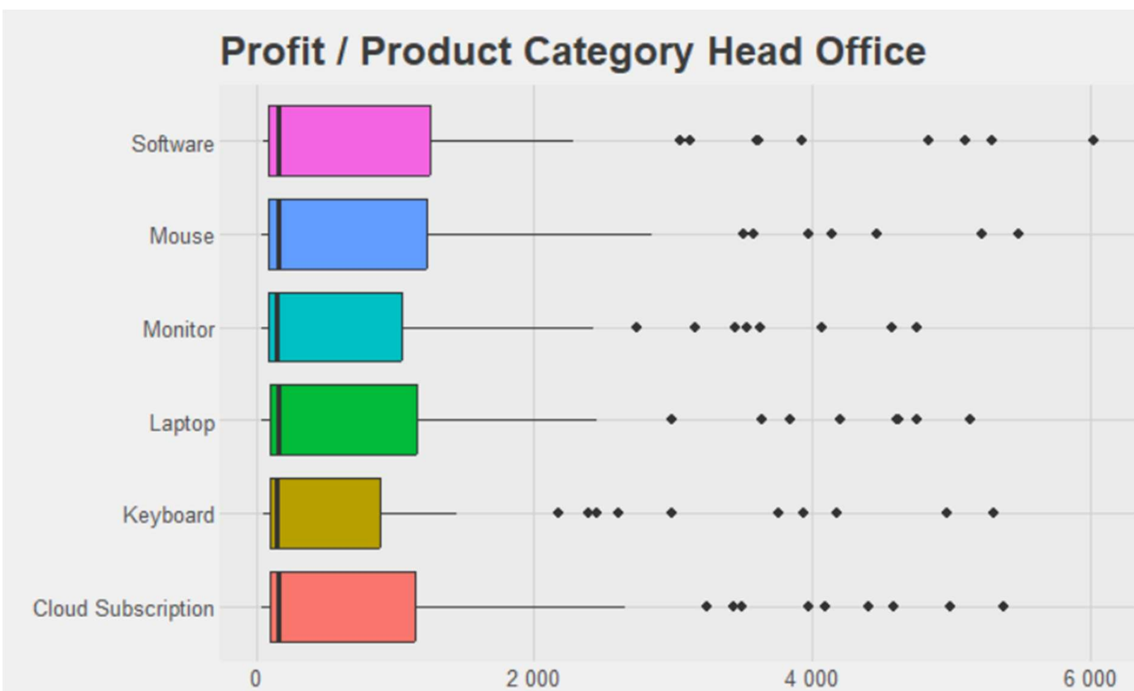

Figure 16: Profit per product category in head office

Monthly sales trends

The graph in Figure 17 is a graphical representation of how the volume of sales changes over time. The monthly sales trend illustrates how total sales volumes change over time. There is a significant dip from October 2022 to December, which gradually increases again from January 2023 to March. This could indicate that the December holiday season negatively impacts sales, as many offices are closed, and it is when most people take leave. The peaks should reflect seasonal demands that may be relevant for the company when planning a marketing strategy.

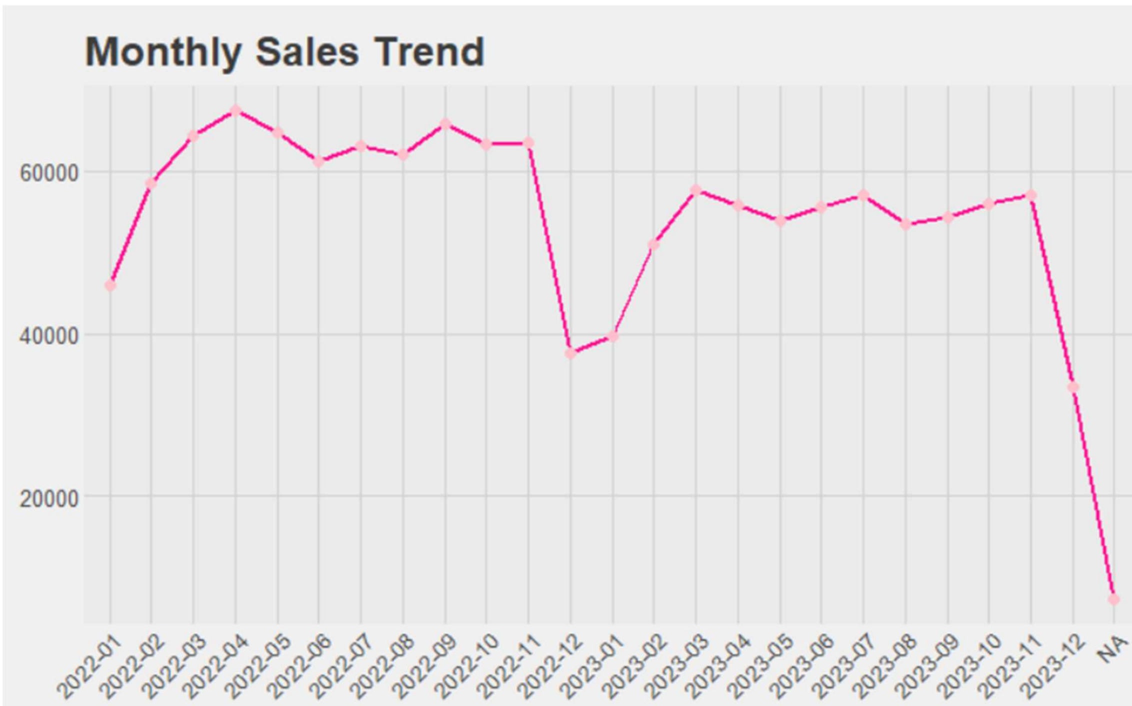

Figure 17: Monthly Sales Trends

Conclusion

Overall, the product analysis assisted the company by revealing insightful conclusions. The head office had a much more standardised method for markups and pricing. Sales were affected by the seasonal changes, meaning that strategic planning should be conducted around holiday seasons. Altogether, these results show that pricing that is done internally is more deliberate, as it experiences less variability than the market dynamics, meaning that the company could take advantage of variability within the market.

Statistical process control

The 2026 and 2027 future sales data were provided in an Excel document. The categories in the document included: the customer and product ID; quantity; order time, day, month, year; and picking and delivery hours. Statistical process control techniques were used in R to monitor delivery times and stability of the various products that included: software, mouse, keyboard, monitor, and cloud subscription.

Initial control charts

The first 30 subgroups of every respective product category were used to create the initial control charts. The standard deviation and mean of each category were calculated to assist in establishing the various control limits.

The rising levels of deviation from the mean were represented by calculating the one-sigma, two-sigma, and three-sigma control limits (LCL and UCL). The one- and two-sigma control limits were used to investigate trends in the data and get an indication of any instability present in the data. The three-sigma control limit was used as a means of defining the primary control boundaries.

To identify if the process functions in statistical control or if corrective action is necessary for variation, the following charts should be analysed:

Cloud subscription

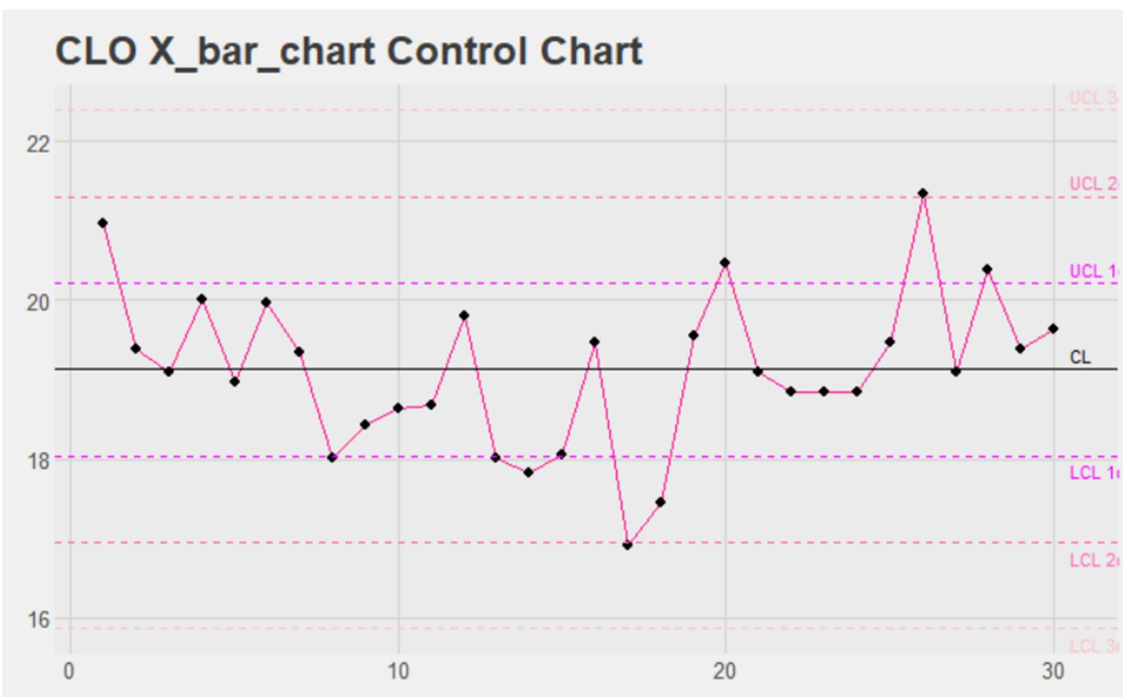

Figure 18: Cloud subscription X bar initial control chart

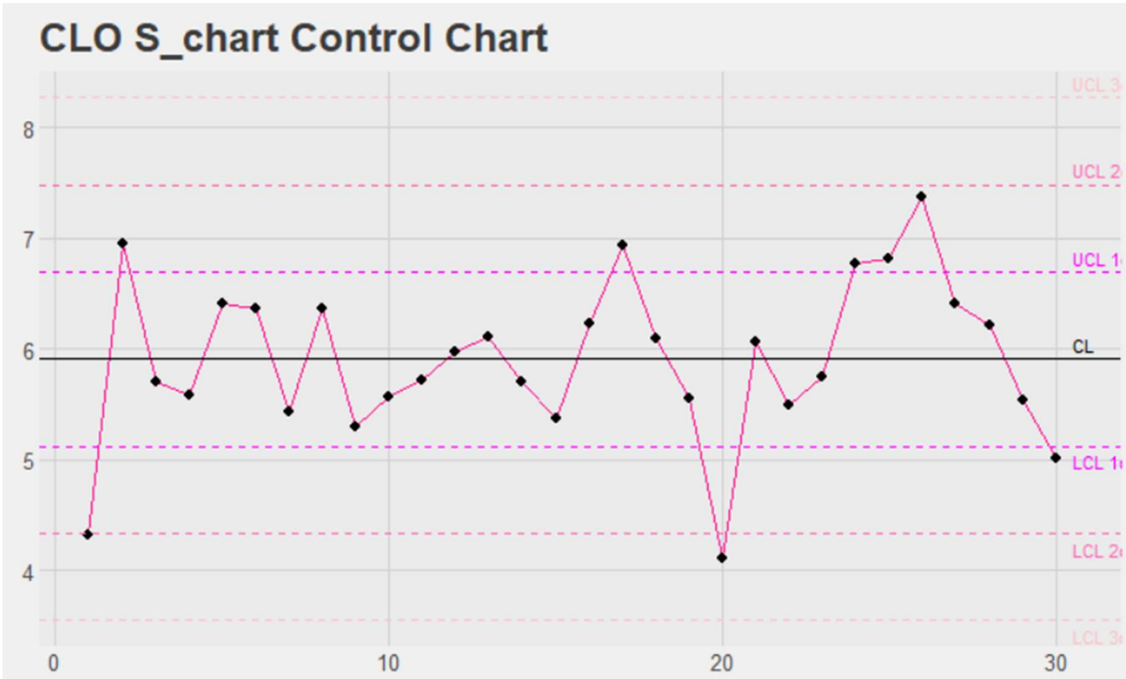

Figure 19: Cloud subscription S bar initial control chart

Laptop

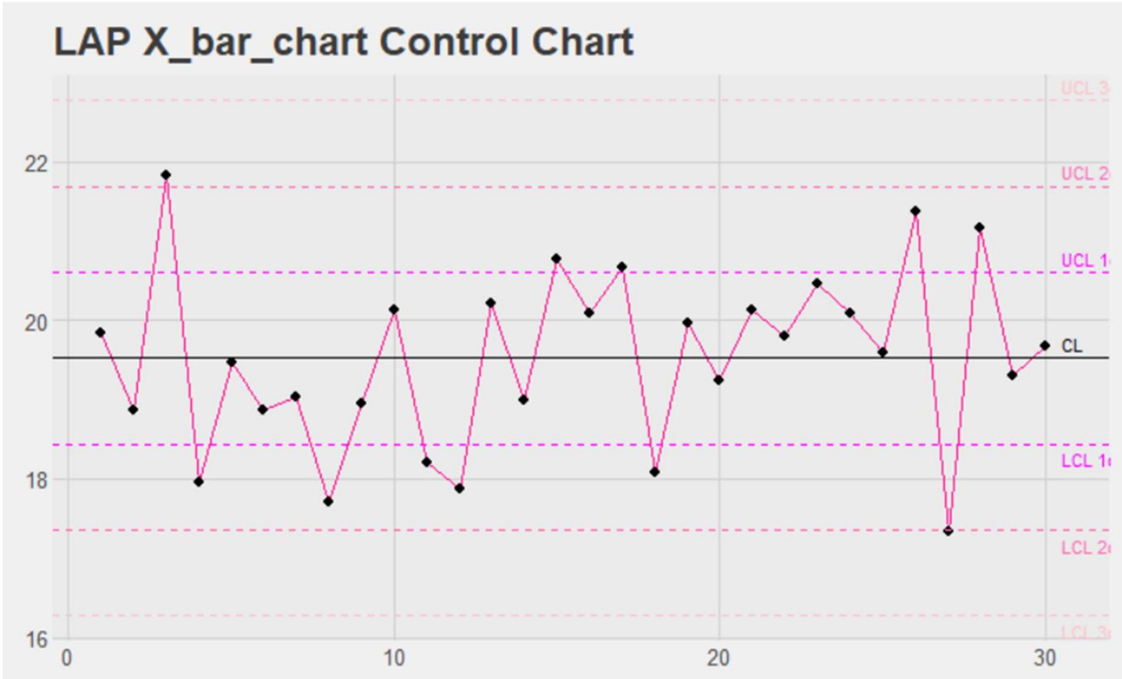

Figure 20: Laptop X bar initial control chart

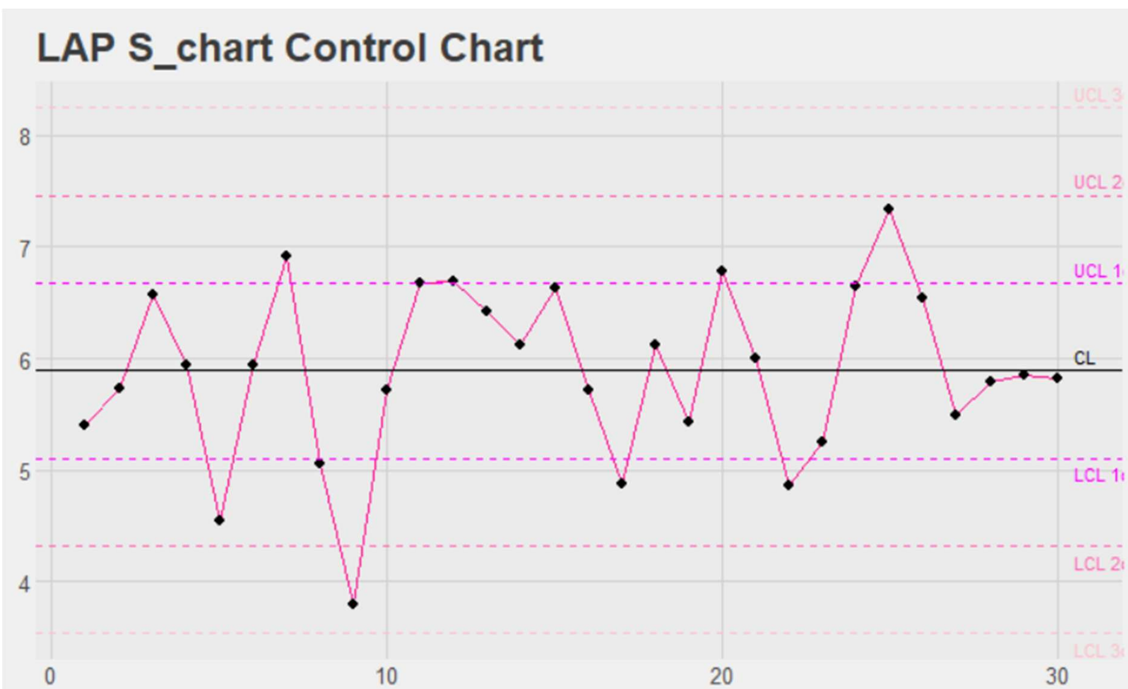

Figure 21: Laptop S bar initial control chart

Keyboard

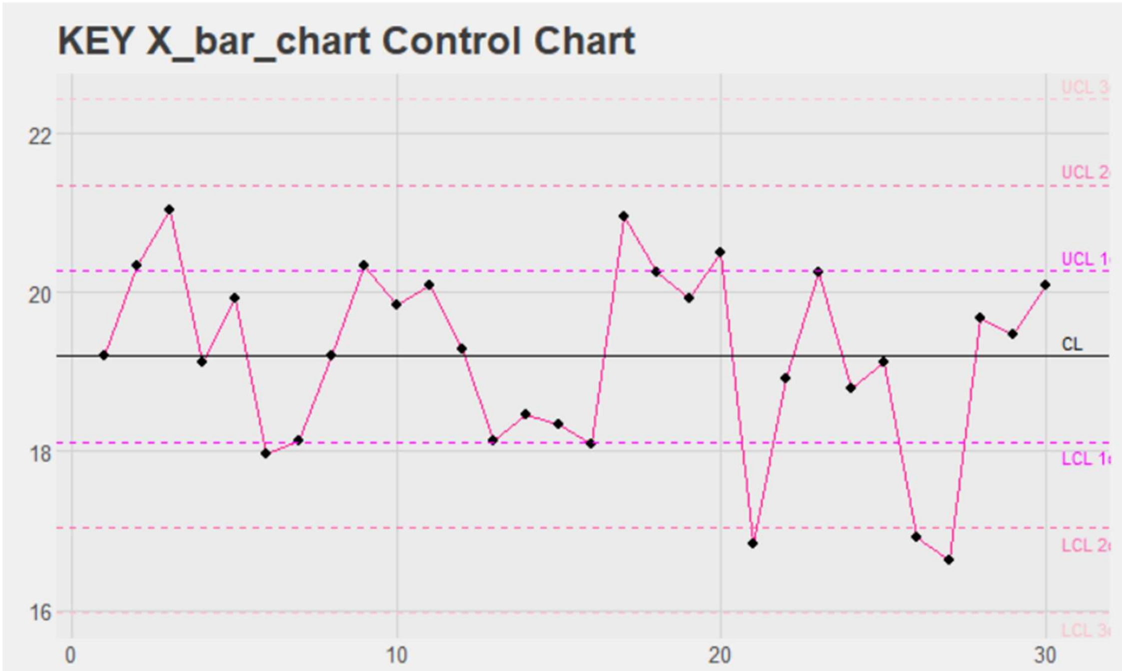

Figure 22: Keyboard X bar initial control chart

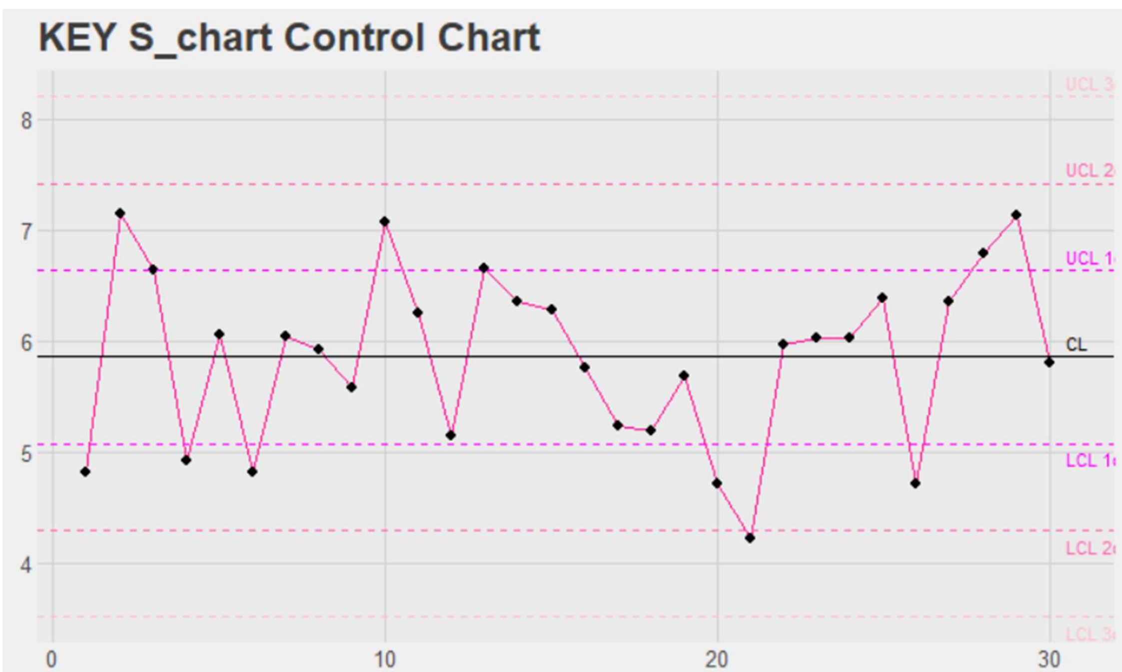

Figure 23: Keyboard S bar initial control chart

Monitor

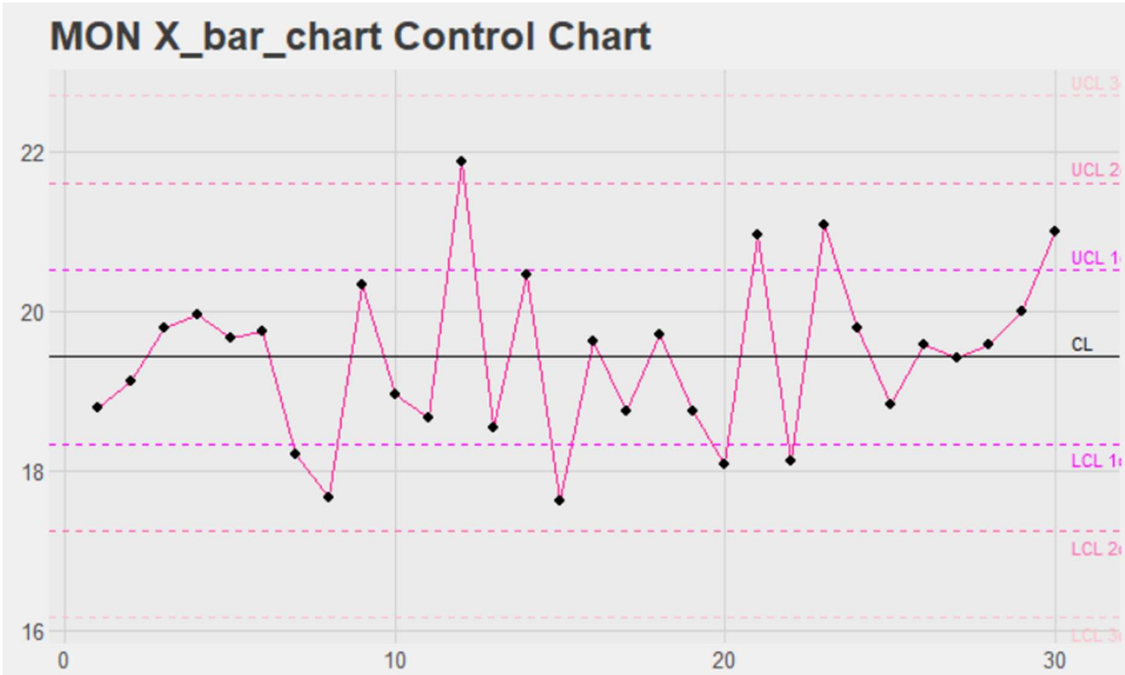

Figure 24: Monitor X bar initial control chart

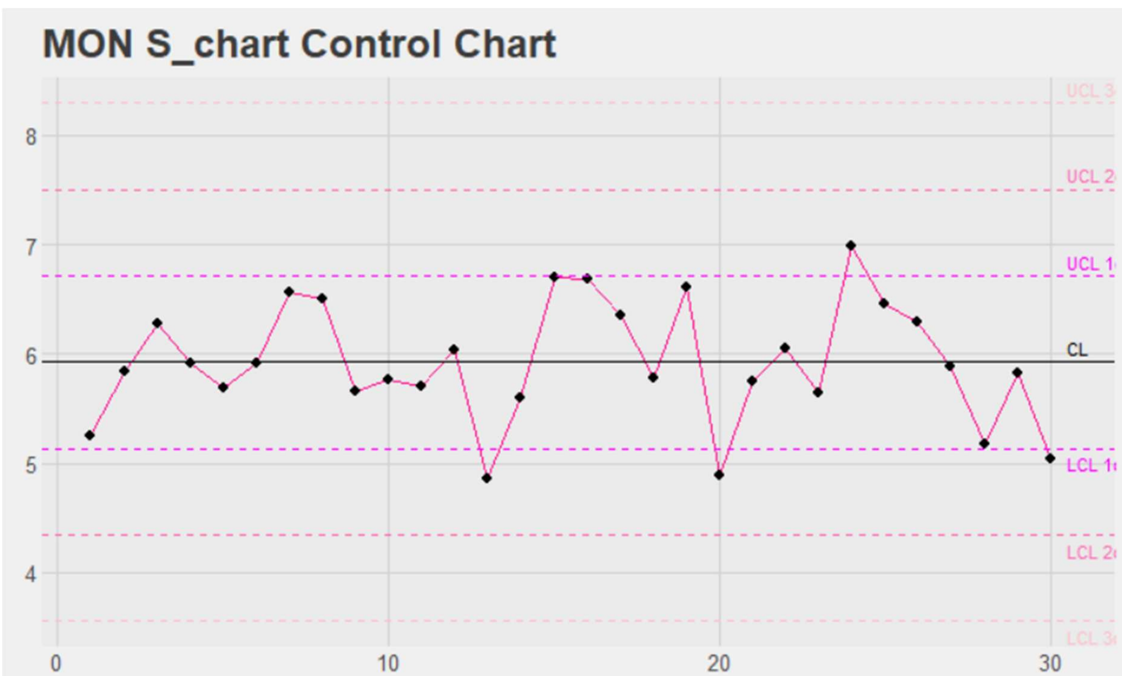

Figure 25: Monitor S bar initial control chart

Mouse

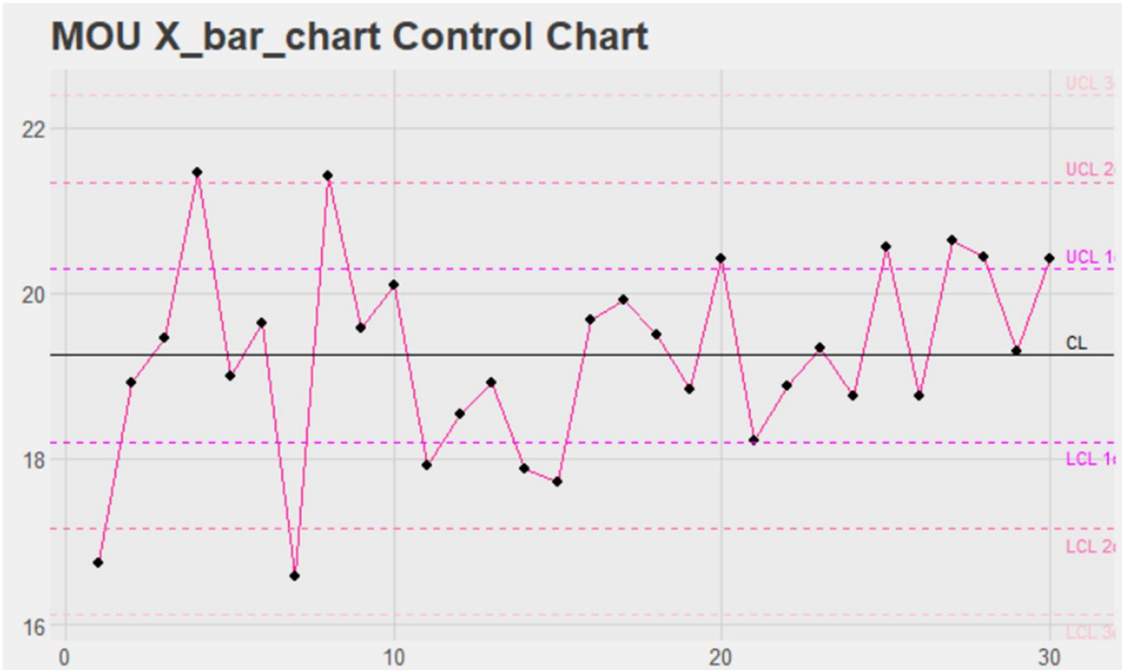

Figure 26: Mouse X bar initial control chart

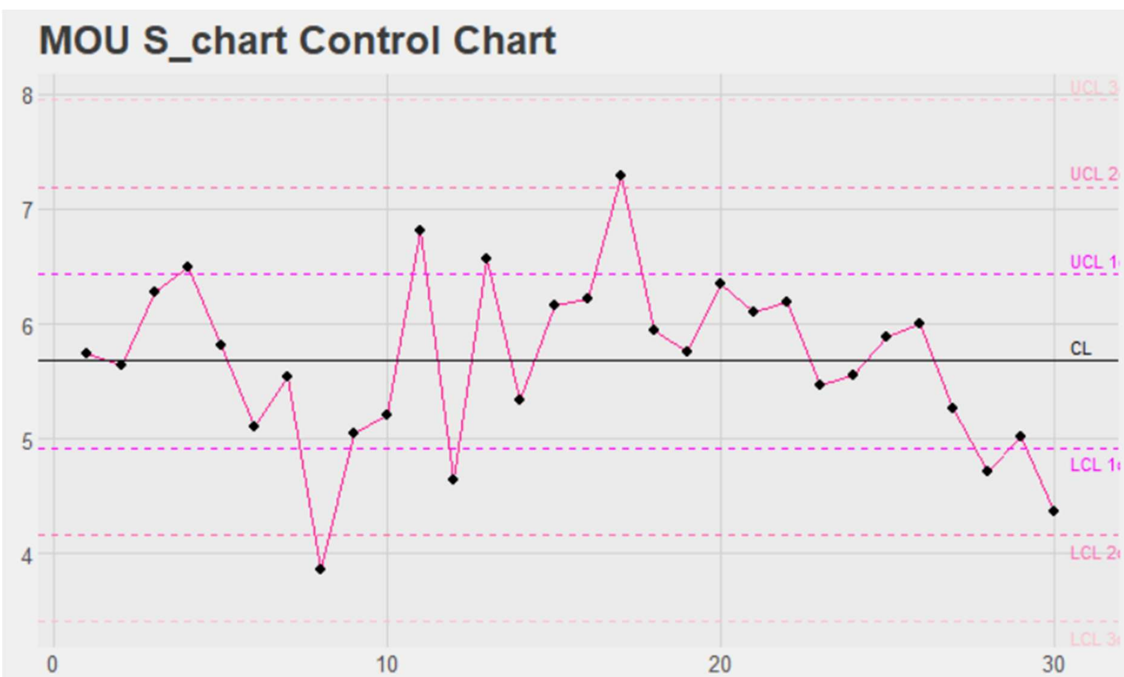

Figure 27: Mouse S bar initial control chart

Software

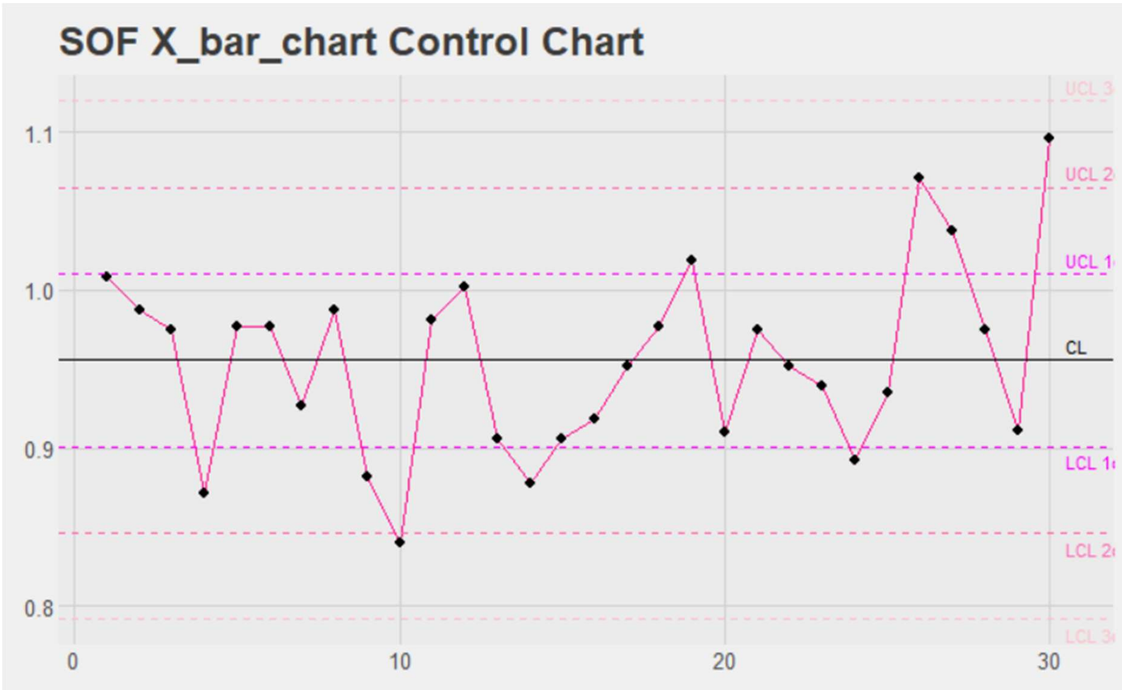

Figure 28: Software X bar initial control chart

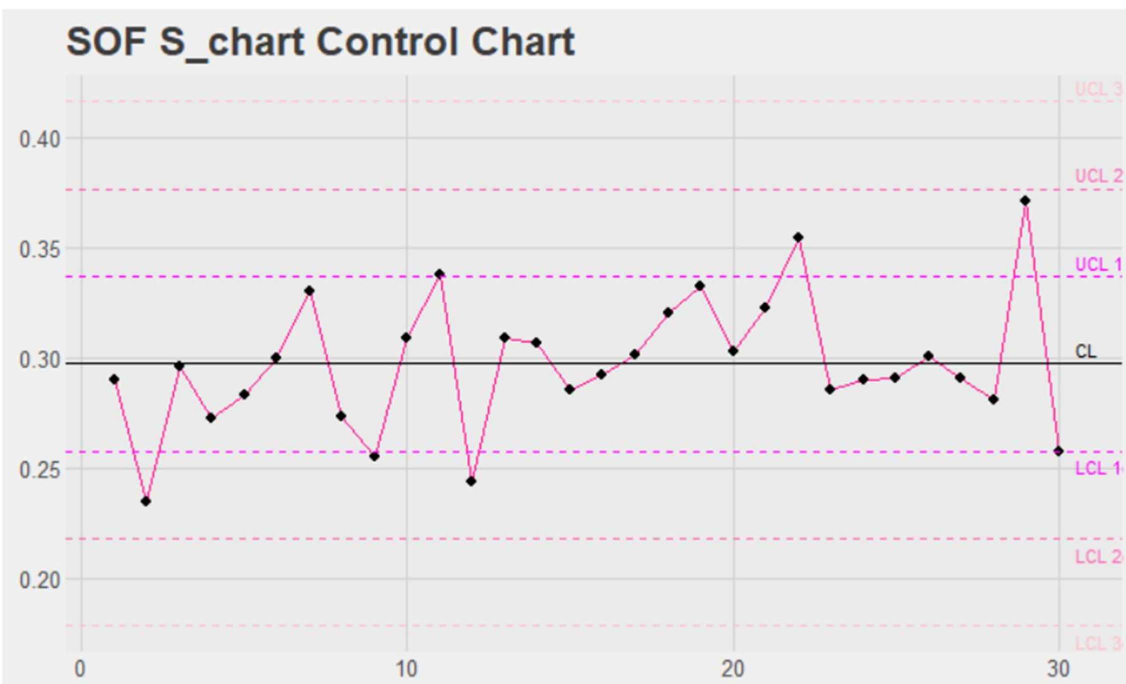

Figure 29: Software S bar initial control chart

Complete control charts

Following the initial control charts, groups of 24 samples were added consecutively onto the control charts until all remaining data had been included in the charts.

From these charts, it is evident that the X bar graphs show a strong upward trend for all product categories.

Cloud subscription

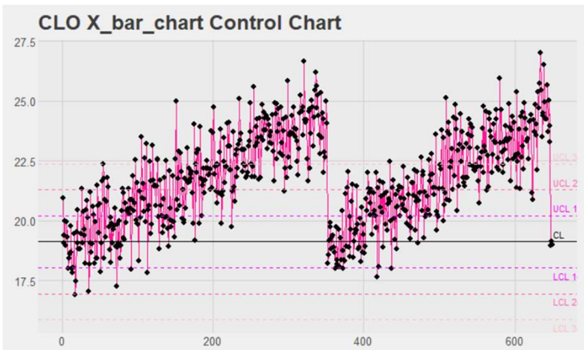

Figure 30: Cloud subscription X bar control chart

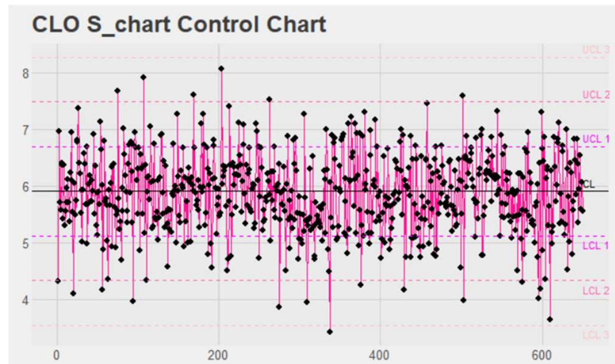

Figure 31: Cloud subscription S bar control chart

Initially, for the cloud subscription, there were various samples that the X-bar exceeded the upper 3-sigma control limit. This was as a result of the upward shift from the baseline. These out-of-bounds samples were an indication of a temporary lack of control with regard to delivery times. The process did, however, reinitiate after sample 351. The upward trend reoccurred, and from sample 490, the samples exceeded the 3-sigma control limit again. There was a notable variation in the cloud subscription delivery times. This could have been a result of fluctuating sales that were influenced by seasonal demands.

Laptop

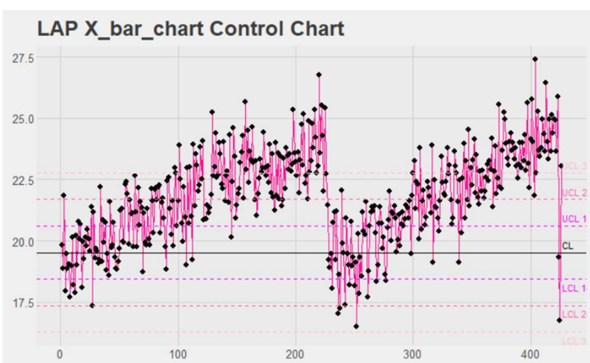

Figure 31: Laptop X bar control chart

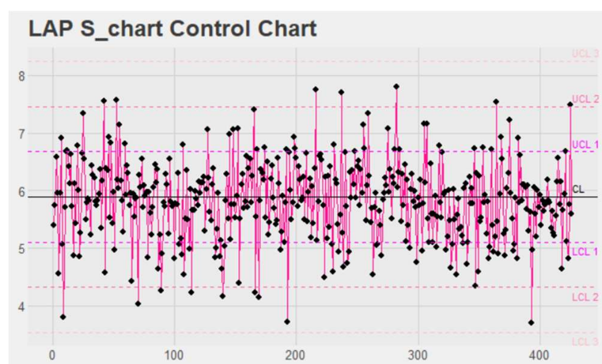

Figure 32: Laptop S bar control chart

Similarly to the cloud subscription, the laptop deliveries also experienced an upward drift. From sample 101, the samples started to exceed the upper 3-sigma control limit. The out-of-control behaviour was recovered from sample 225 until it reoccurred from sample 301. Since the standard deviation managed to remain within the 3-sigma control limit, the variation present was random.

Keyboard

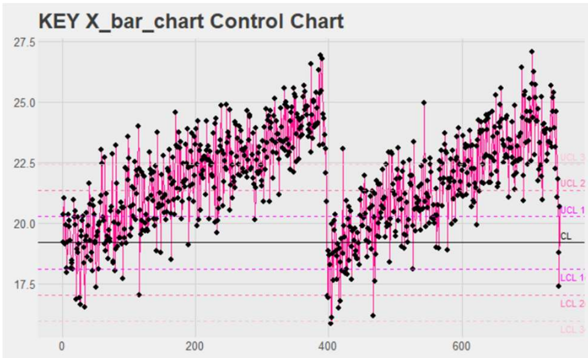

Figure 33: Keyboard X bar control chart

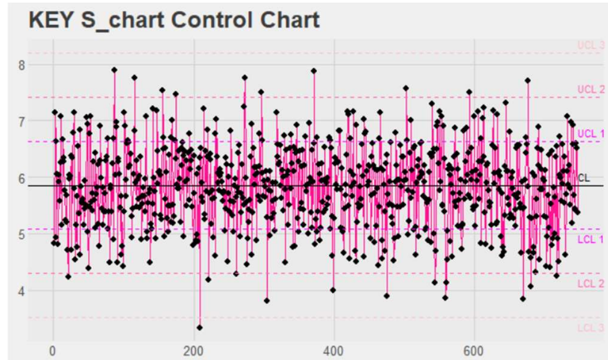

Figure 34: Keyboard S bar control chart

Similar variation was experienced for the keyboard delivery times as in the cloud subscription and laptop. The out-of-bounds samples began at sample 59, where deliveries started to take longer. They recovered at sample 396, then proceeded to upwardly head out of bounds from sample 490.

Monitor

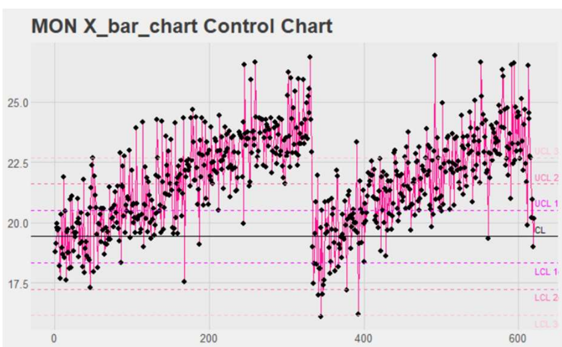

Figure 35: Monitor X bar control chart

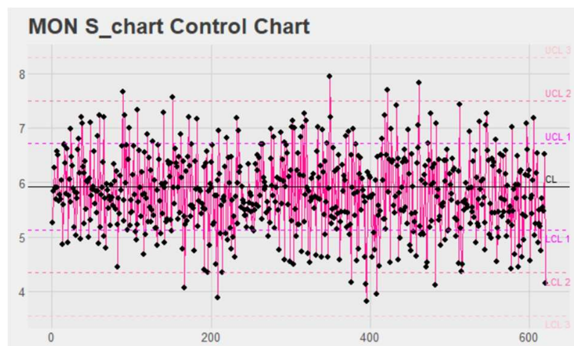

Figure 36: Monitor S bar control chart

For the monitors, the first out-of-bounds samples started at sample 106. Temporary control was gained, but the upward trend continued. There were low variations between the samples, as observed in the s chart.

Mouse

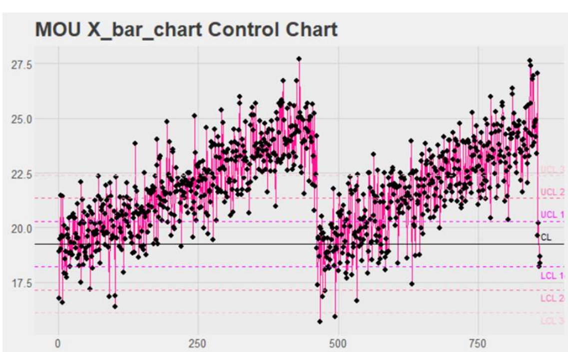

Figure 37: Mouse X bar control chart

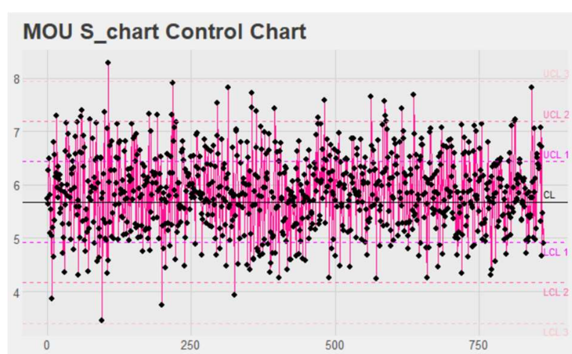

Figure 38: Mouse S bar control chart

Delays in delivery times started appearing from sample 137. This was recovered from sample 461, but the delays once again continued from 546. The deliveries were generally fairly consistent, with the occasional delivery time spikes

Software

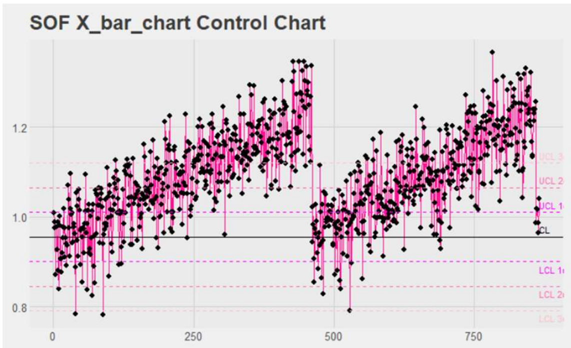

Figure 39: Software X bar control chart

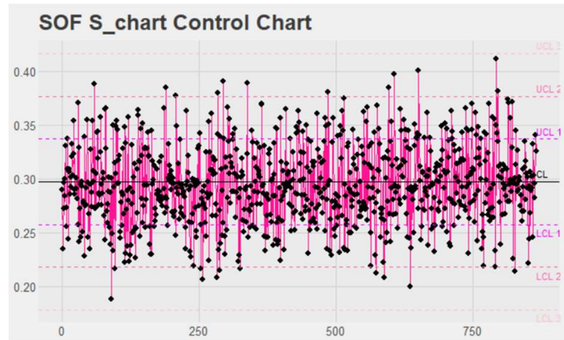

Figure 40: Software S bar control chart

There is a low product variation for the software category. The delivery times followed the same trends as the other product categories, where an upward increase eventually surpasses the UCL3 line at sample 162. This was then recalibrated at sample 460. As before, the upward trend increased until sample 570 passed the UCL3 line.

Conclusion

The delivery times of all categories increased over the period of analysis, as indicated by the gradual increase observed in the X-bar charts. They all would regain statistical control at some point, but then drifted above the three sigma limit again. Since this was true for all product categories, there is some factor that causes these periods of no statistical control, which should be investigated by the company, as remedying it would likely positively influence all product categories. However, since the monitor and software categories had the lowest levels of variation among the products, they had more stable delivery processes. Statistical process control should continue to be undertaken in the future to assist the company in reducing delivery time delays.

Process Capabilities

Cp, Cpl, Cpk, and Cpu statistics were used to conduct a process capability investigation for each product category. This is used to determine whether or not the product will be delivered within the appropriate specification limits. LSL is 0 hours, and USL is 32 hours. Since delivery times cannot be negative, the Cpu will provide more relevant information than the CPL.

Most products have Cpk, Cpu, and Cp values lower than 1. This poorly reflects their ability to comply with the delivery time limits, meaning that most deliveries have taken more than 32 hours. In reality, this will have extremely negative impacts on the business as the customer dissatisfaction will increase due to poor service levels.

The very small Cpk variation across the products indicates that the reason for the lack of control is systemic and not related to specific products. The issue could be found within shared logistic processes.

The Cpl values are almost always higher than the Cpu values. This shows that late deliveries are the major issue that the company is facing, as opposed to early deliveries.

However, the software products have Cpu and Cp values that are very high, and the Cpk is above 1. This means that this category has managed to stay within the target limits. This is a logical deduction, as software product deliveries are instantaneous, unlike the physical delivery of physical products. Unexpected issues such as transportation delays are not a risk for this category.

Process capabilities ofr each product category

	Product	Cp	Cpu	Cpl	Cpk
CLO_CP	CLO	0.8977458	0.7167378	1.078754	0.7167378
LAP_CP	LAP	0.8950269	0.6928906	1.097163	0.6928906
KEY_CP	KEY	0.9171375	0.7293536	1.104921	0.7293536
MON_CP	MON	0.8890490	0.6995705	1.078528	0.6995705
MOU_CP	MOU	0.9151848	0.7265710	1.103799	0.7265710
SOF_CP	SOF	18.1352369	35.1876018	1.082872	1.0828720

Figure 41: Process capabilities

Conclusion

The investigation of the process capabilities uncovered that the company has a lack of process control. This could be negatively affecting the company as it is decreasing the service level. The software category was the only one that met the capability standards. This shows the divide between physical and digital product distribution. The company needs to make systematic changes to improve the logistics and delivery time.

Process control

The process control issues for the six product categories were uncovered through the s- and x-bar charts. The three rules below were used as a basis:

Rule A

1 s sample outside of the upper +3 sigma-control limits for all product types:

Product	**Outlier details**
Cloud subscription	None identified
Laptop	None identified
Keyboard	None identified
Monitor	None identified
Mouse	Sample 107, Value=8.285551, above UCL3
Software	None identified

Rule B

Find the most consecutive samples of s between the -1 and +1 sigma-control limits for all product types:

Product	**Start Position**	**End Position**	**Longest consec. sample within range**
Cloud subscription	477	498	22
Laptop	395	417	23
Keyboard	652	666	15
Monitor	236	253	18
Mouse	455	466	12
Software	829	850	22

Rule C

4 consecutive X-bar samples outside of the upper, second control limits for all product types:

Product	**Number of instances**	**Example Ranges**
Cloud subscription	15	(120-123), (165-172), (177-181), (554-562), (564-622),(625-647)
Laptop	10	(115–121), (128–139), (151–164), (349–355), (357–367), (369–423)
Keyboard	27	(97–100), (170–173), (175–178), (671–692), (694–720), (722–742)
Monitor	23	(132–135), (169–175), (177–184), (562–589), (591–610), (612–616)
Mouse	29	(197–200), (209–216), (221–225), (784–797), (799–803), (805–856)

Software	30	(129–133), (198–202), (204–208), (758–794), (796–835), (837–859)

Conclusion

This analysis shows that most of the product categories were within the control limits. However, small deviations that were continuously seen indicate that there is process instability across the different categories. If stability is short-term, it is somewhat acceptable. What is not acceptable is the long-term process control issues that are seen in the software, keyboard, and mouse categories. These issues stem from systematic issues within the company, rather than specific product issues. These issues should be uncovered for the company to be able to make necessary improvements.

Risk, data correction, and optimising for maximum profit

Type I error

Rule	Probability of Type I error
A	0.001349898
B	0.6826895
C	2.678772e-07

Type II error

Probability= 0.8411783

Data Correction

Product

By referring to Figure 41, it can be seen that the data correction effectively corrected the output of prices per category. This visual is much more realistic than that of Figure 9. Laptops are seen to be much more expensive than the other categories. The value spread is also the largest, as different types of laptops will have high variations in price depending on the complexity of the product. There is much more opportunity for variation within the laptop category than there would be in other categories, such as the mouse, as the mouse is limited in how complex it can get, so all mice will be of similar selling price. Following the laptop is the monitor, which is also more expensive than the other product categories. Once again, this is what the expected distribution should look like. In the real world, laptops and monitors are much more expensive than cloud subscriptions, mice, keyboards, and software.

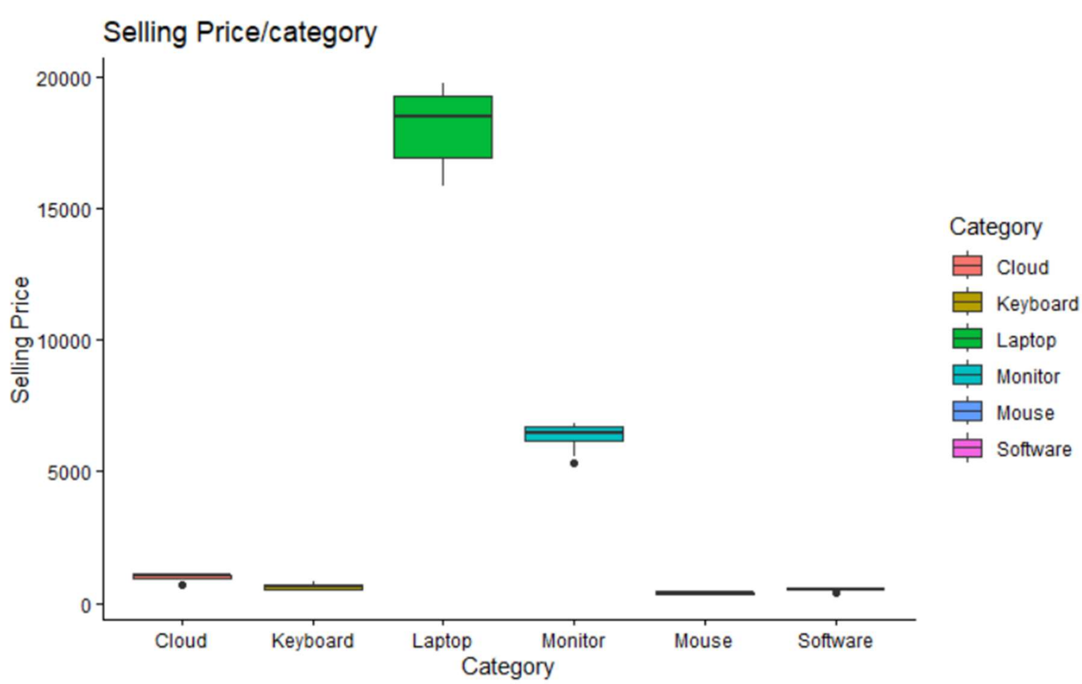

Figure 41: Selling price per product category

Figure 41 is further supported by Figure 42, which also visually indicates that the laptop category has the largest selling price, which is followed by the monitor category. The cheapest category is the mouse. The distribution in Figure 42 is that of an exponential distribution, unlike the more uniform distribution that was seen in Figure 7.

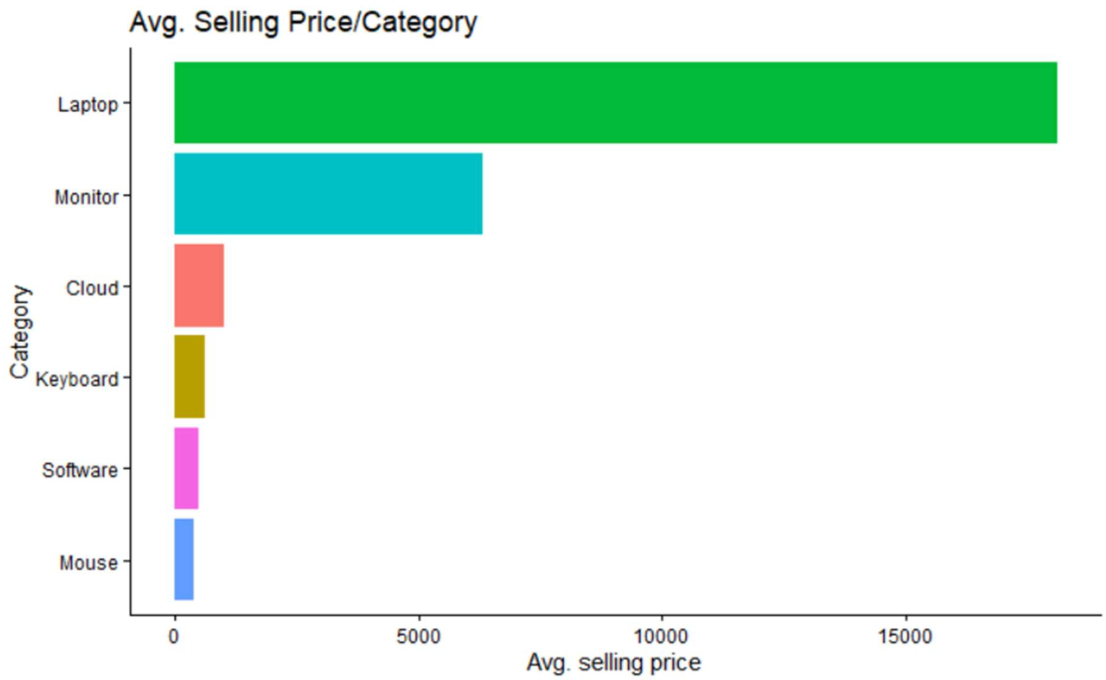

Figure 42: Average selling price of each product category

The data correction procedure that aimed to correct product categories did not affect the markup distributions.

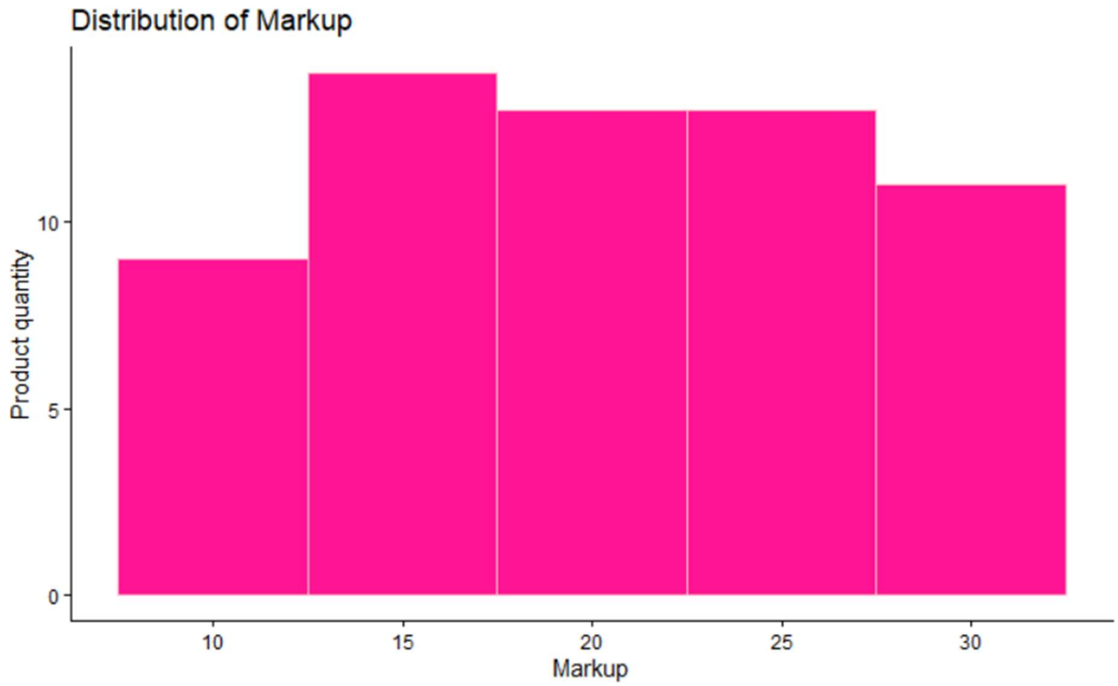

Figure 43: Distribution of markup

However, the distribution of markups for the various product categories did alter significantly. This can be seen in Figure 44. Monitors have the largest range of markups after the correction. The software category has the lowest markup.

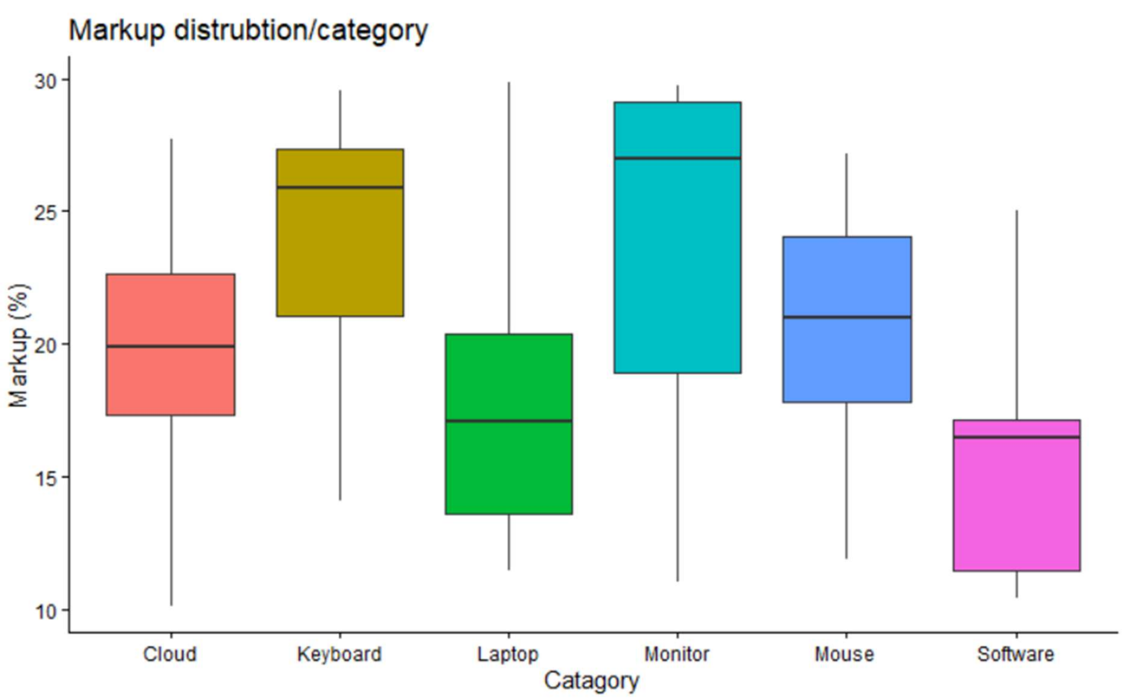

Figure 44: Distribution of markup per product category

In correlation with the changes in selling prices and markups, the profits per product also changed dramatically. Naturally, the laptop category experienced the highest profit distribution as seen in Figure 45. The profit distributions are directly proportional to the selling prices.

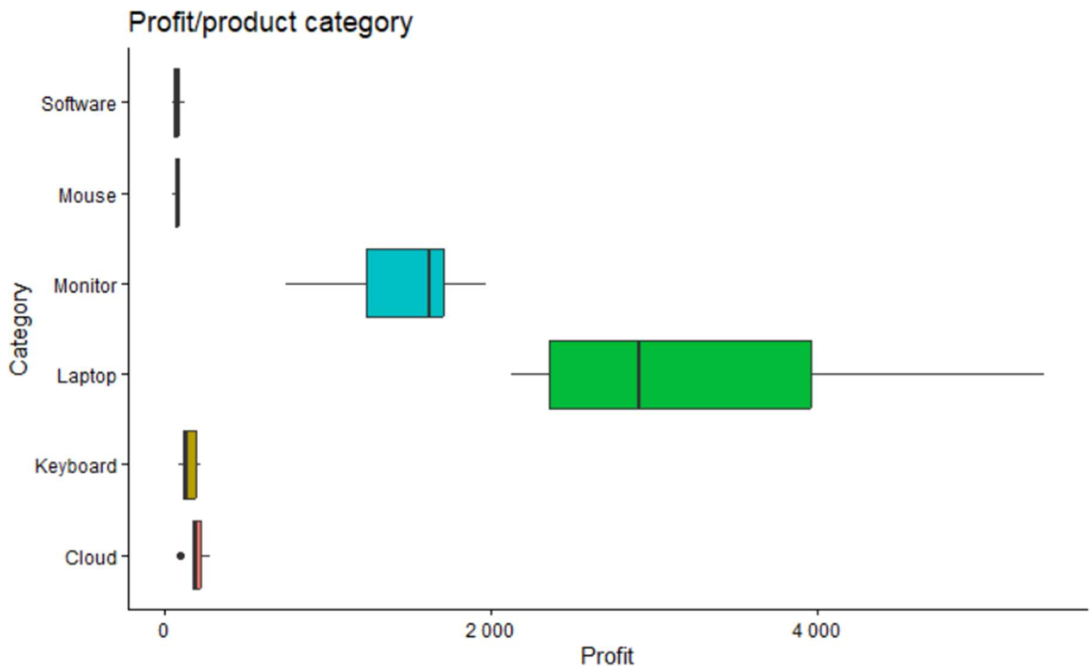

Figure 45: Profit per product category

Head Office

Similar to the change in selling prices seen in the products data, the products' head office also saw a massive shift. The distribution in Figure 46 is almost identical to that of Figure 41. The laptop category is seen to have the largest selling price and the widest spread of prices, as in Figure 41.

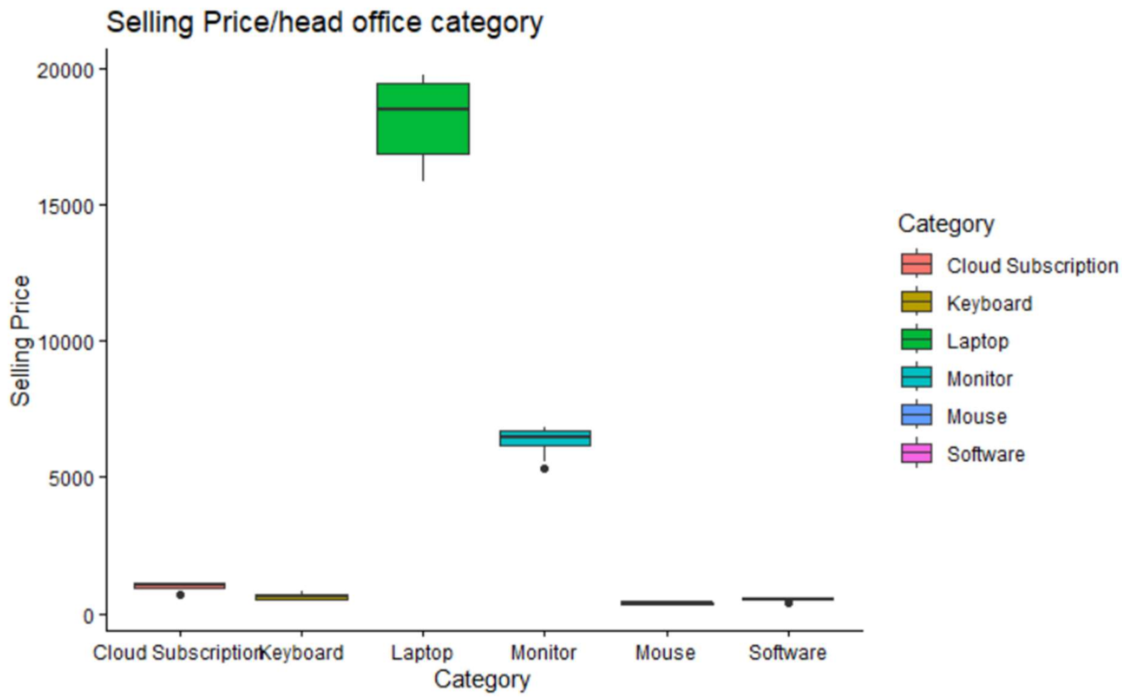

Figure 46: Selling price per product category (head office)

The same can be said for the spread of average selling prices in Figure 47. This distribution is the same as that in Figure 42, and supports the information seen in Figure 46.

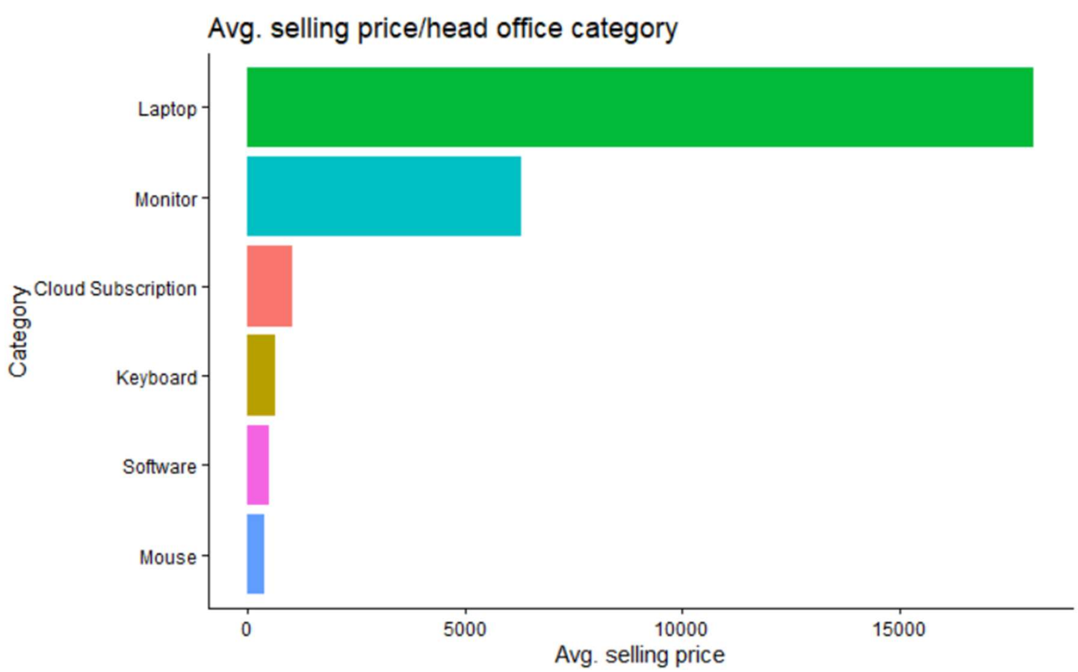

Figure 47: Average selling price per product category (head office)

Similarly, the distribution of markups per product category in the head office in Figure 48 is almost identical to that seen in Figure 44. With the monitor category having the largest distribution and the software having the smallest.

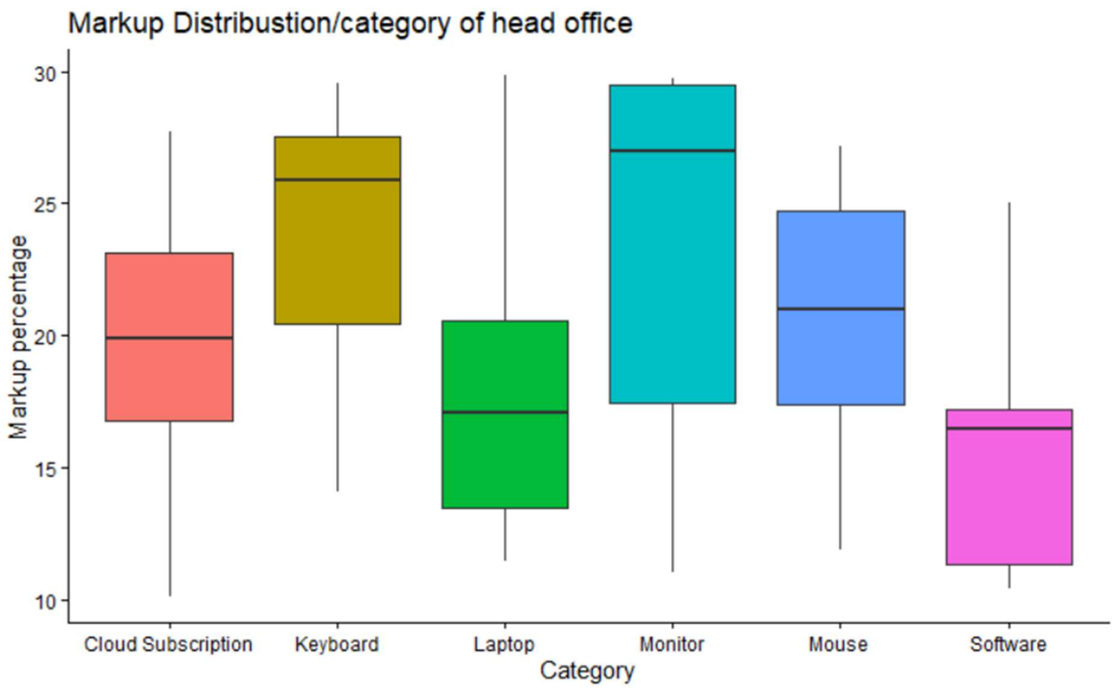

Figure 48: Distribution of markups of product categories (head office)

The profit is directly proportional to the selling prices. In Figure 49, it can be seen that laptops render the highest profit. This is followed by monitors. The profit has significantly changed after data correction.

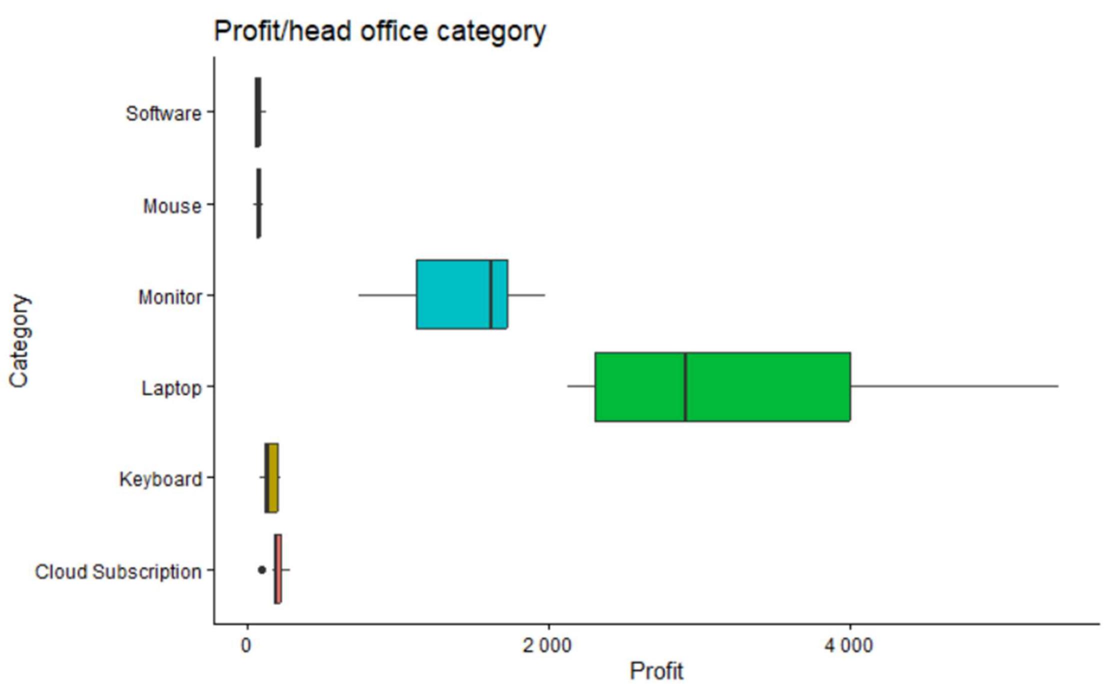

Figure 49: Profit per category (head office)

Sales

Product Category	Mean selling price (R)	Total sales (R)
Keyboard	644.66	5378599
Monitor	6310.53	3773414
Mouse	394.70	86027413
Laptop	18086.43	43126708
Software	506.18	7261887
Cloud	1019.06	4867781

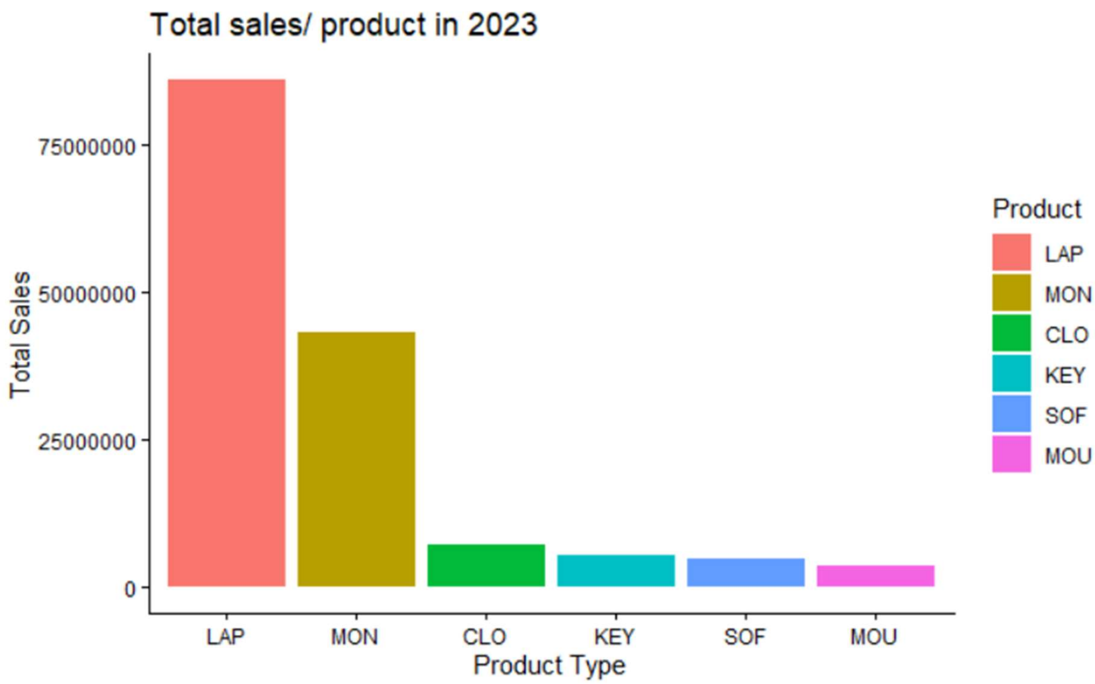

Figure 50: Total sales per product in 2023

The greatest contributor to the overall sales was the laptop. This is a logical conclusion since it had the highest selling price by a long way.

Profit optimisation

Coffee shop 1

The assumptions that needed to be made to perform optimisation were as follows:

- Working hours: 8am-5pm (no break)
- Demand was not capped

By referring to Figure 51, it is evident that every time a barista was added, the waiting time significantly decreased. This can be seen in the exponential decay of the graph. This is further supported by Figure 52, where the number of customers that were attended to linearly increased as the number of baristas increased. They have a directly proportional relationship.

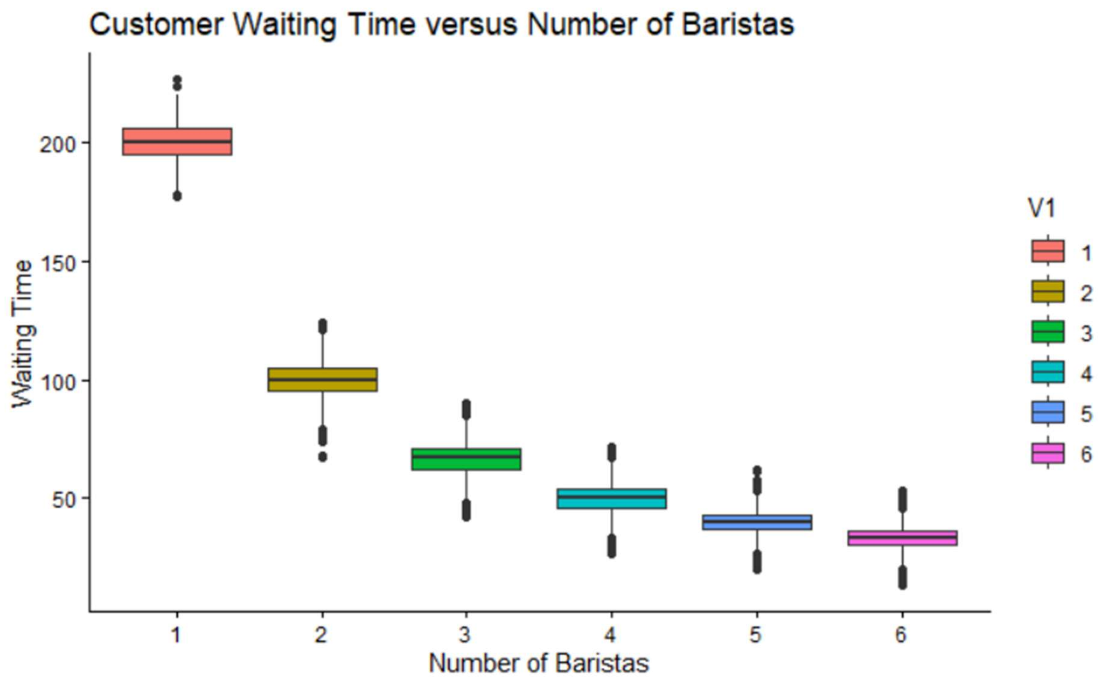

Figure 51: The effect of the number of baristas available on the customer waiting time

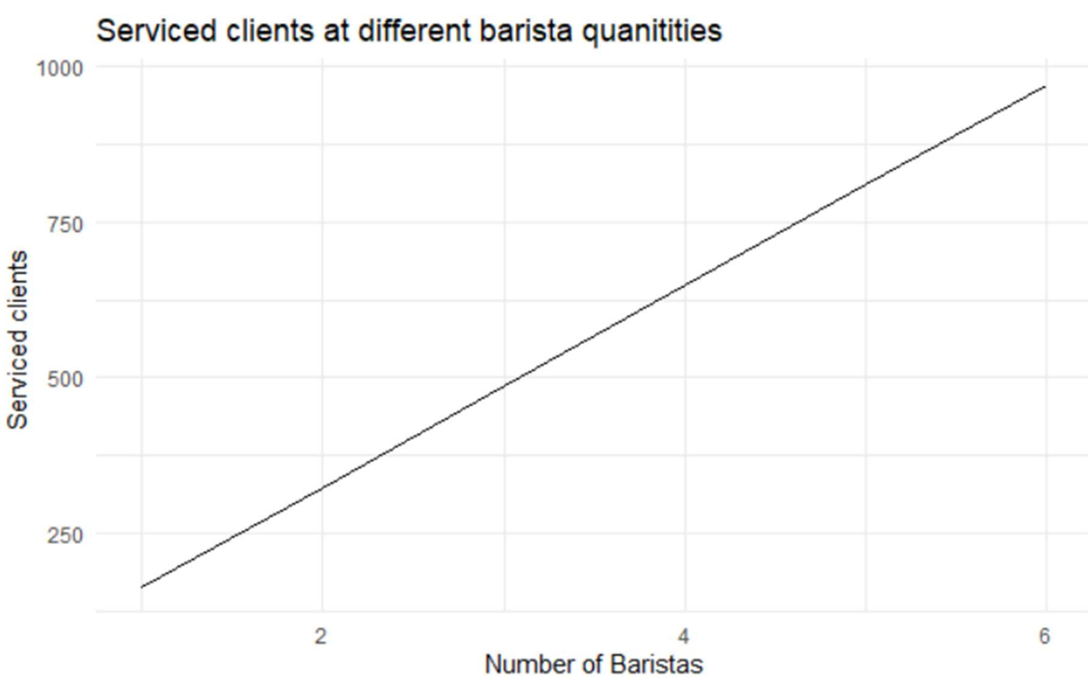

Figure 52: The effect of the number of baristas on the number of clients attended to

The more customers that get attended to, and the less time customers wait, the better the service level and the more profit the company can make. This is seen in Figure 53 as the profit made by the company and the number of baristas have a positive linear relationship. The profit is close to R23000.

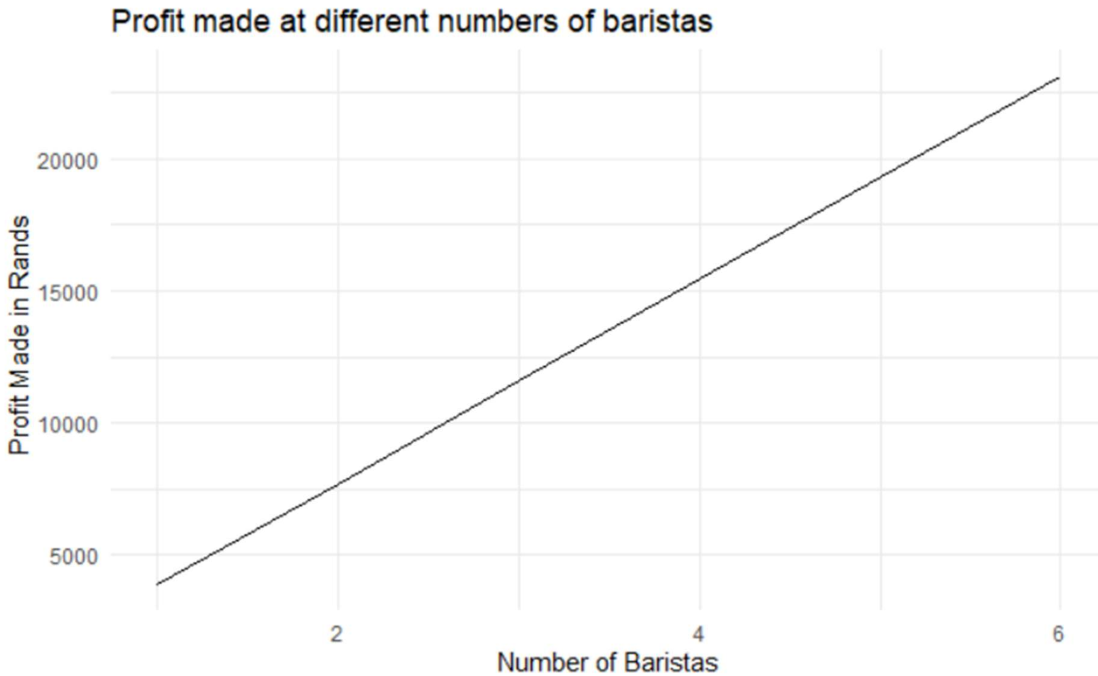

Figure 53: The profit made by the company versus the number of baristas employed

As seen in Figures 53 and 54, the optimal number of baristas to be employed is 6. It is logical that this number is capped, as it still costs the company money to employ baristas. Once the operation capacity has been met, then the labour hours paid for start going to waste.

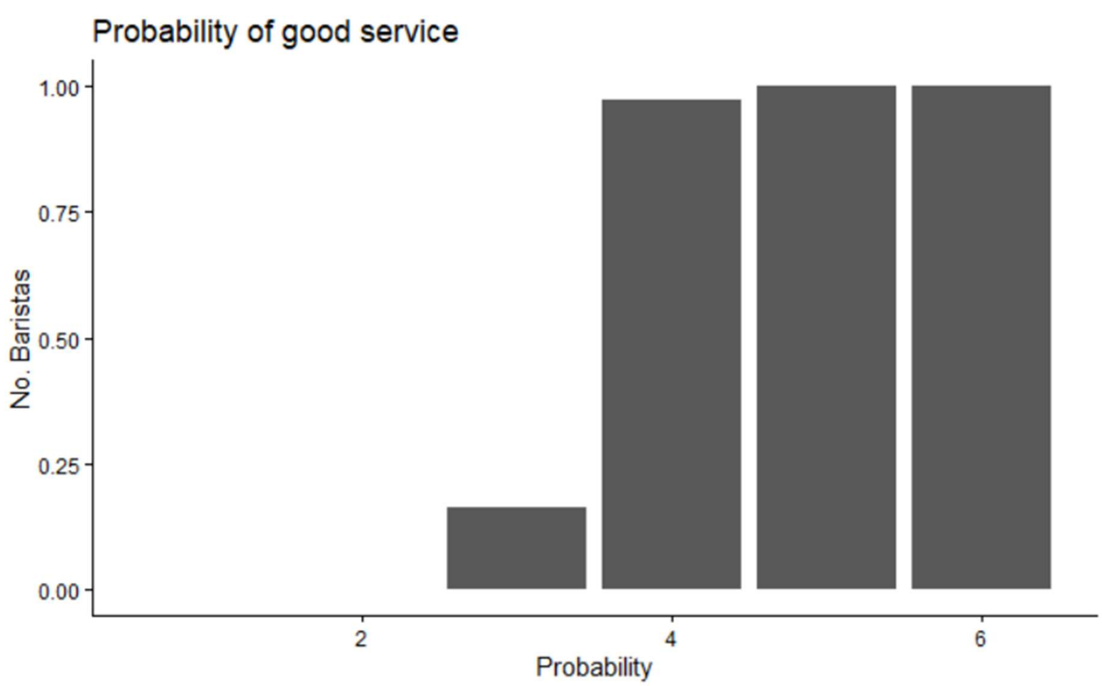

Figure 54: The probability of good service

Taguchi loss function

The Taguchi loss function can be used to quantify the loss caused by processes deviating from the target value. In this case, it could be used to quantify the loss caused by the poor customer satisfaction that results from waiting times being longer than a minute. The loss increases as the waiting times increase, which shows how future sales will decrease due to a poor reputation. This supports the decision to have more baristas to not only increase the current profit but also the profit made in the future, earned from having a good reputation in the marketplace.

Coffee Shop 2

Every barista added results in a decrease in waiting time. This is a logical conclusion, as if there are more employees serving customers, the customers will have to wait less time to be helped, as the load is better distributed. The decrease follows an exponential decline. The difference between the first and the second coffee shop is that the second coffee shop is likely to have less capacity to serve customers, as the waiting time for 6 baristas is much less than the first coffee shop.

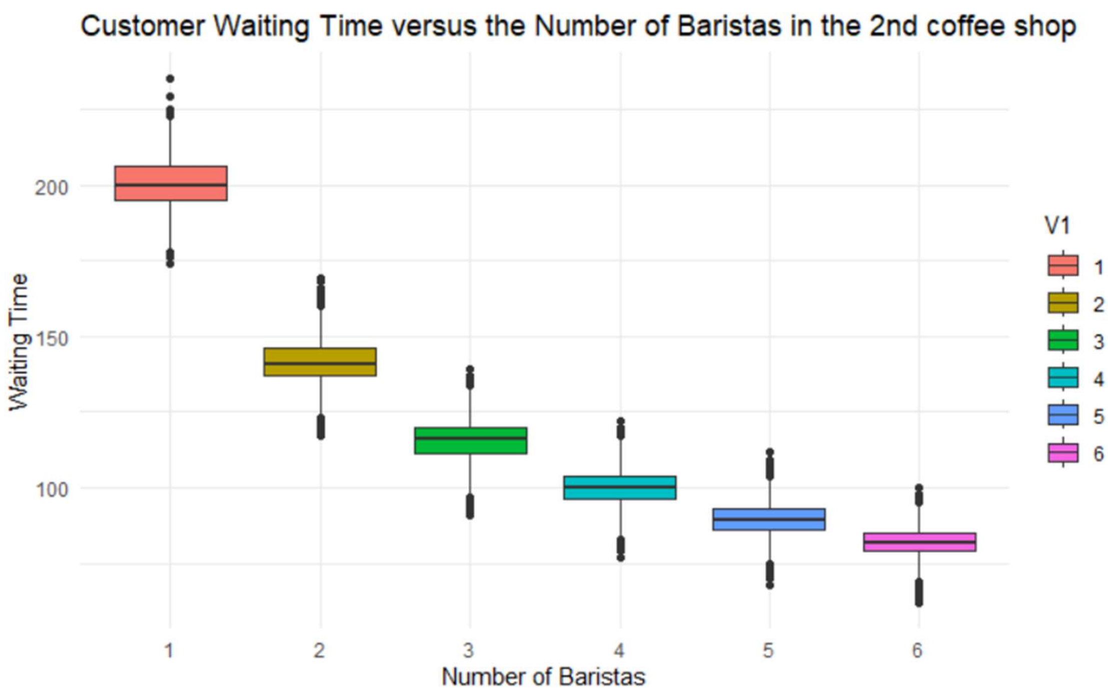

Figure 55: Customer waiting time vs number of baristas

This is further supported by Figure 56, as it can be seen that fewer customers are served. The capacity of this coffee shop is obviously smaller than that of coffee shop 1. This could be caused by a lack of space, or inefficiencies that result in longer waiting times, and in turn fewer customers get served.

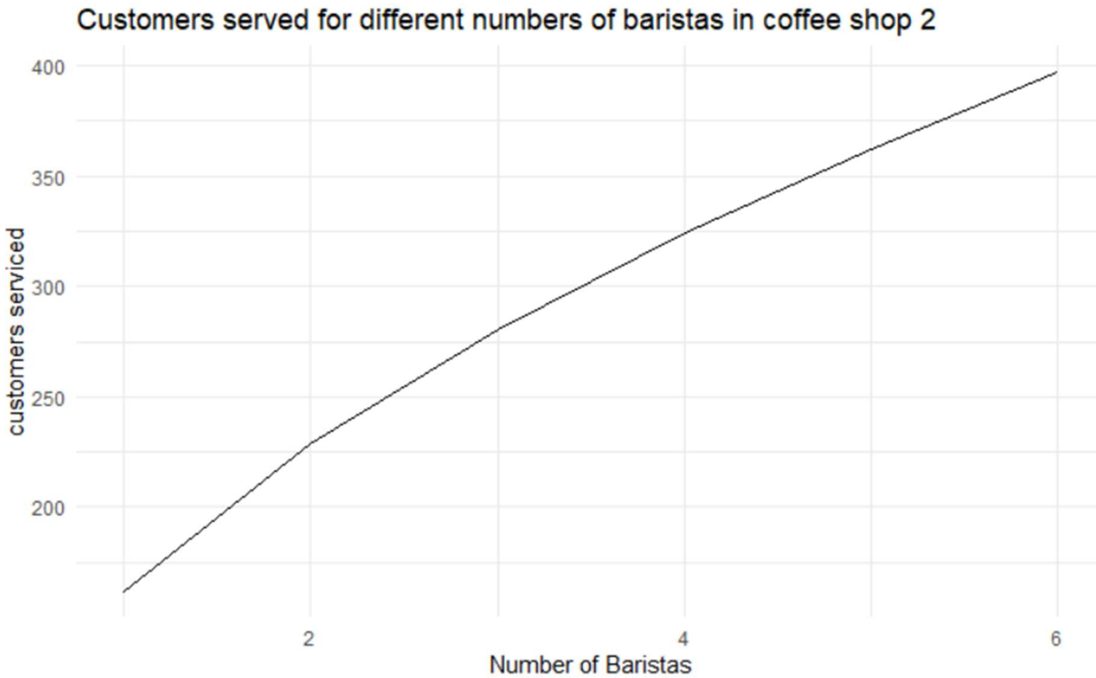

Figure 56: customers served vs number of baristas

As the number of baristas approaches 6, the profit that had been gradually increasing begins to plateau. This means that the ideal number of employees should be 6. This will allow the company to make a profit of R5910, as seen in Figure 57. Employing more baristas will cause an imbalance in the cost-to-profit ratio and end up costing the company more money than they would make. This is also supported by the probability of reliable service, as it is 97.44% for six baristas. This can be seen in Figure 58.

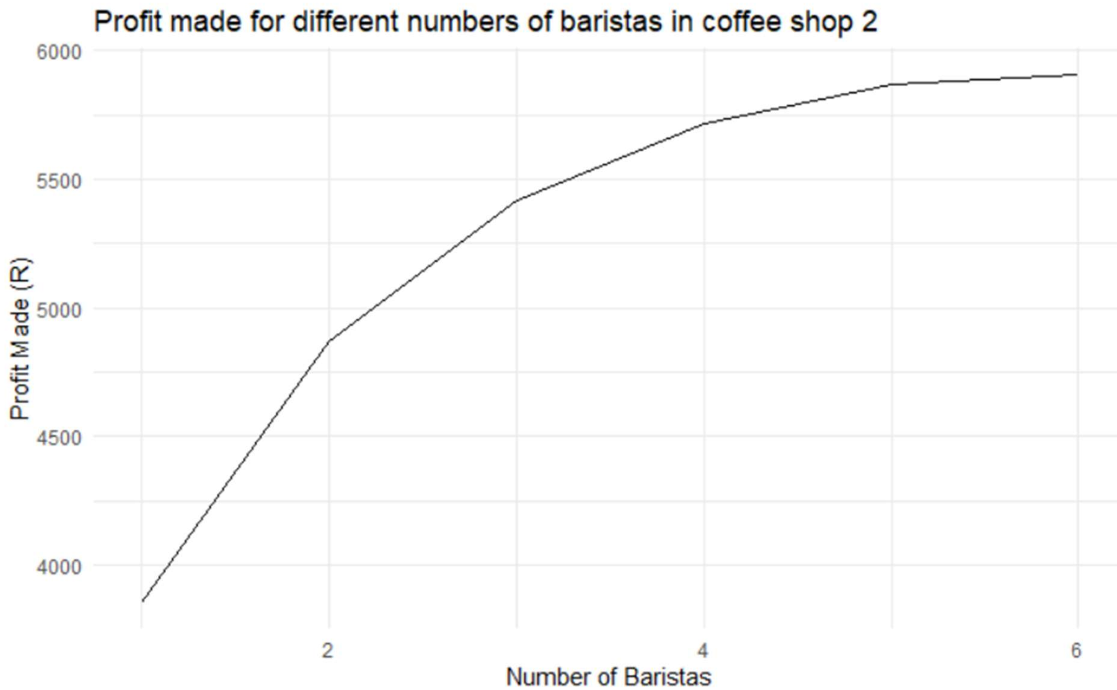

Figure 57: Profit vs number of baristas

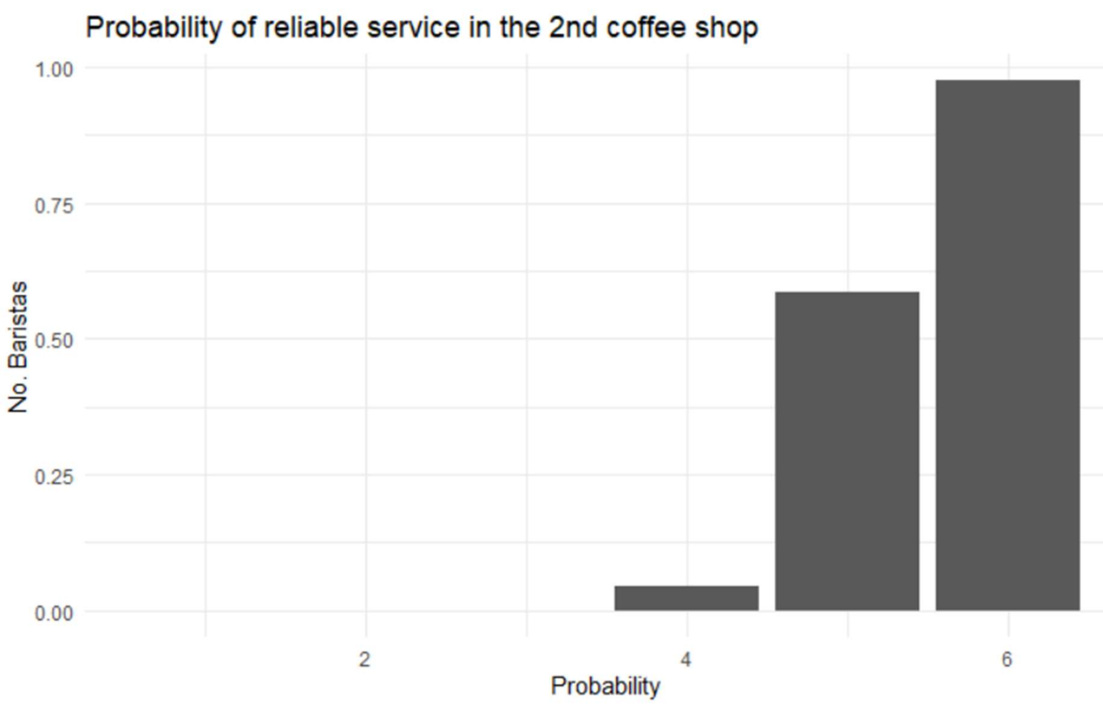

Figure 58: Probability of reliable service

Conclusion

Type 1 errors were highly unlikely in most rules. This is less true for Type 2 errors. The data correction of the selling prices, markup, and profit distributions revealed more realistic deductions. This became clear when looking at the laptop and monitor categories, which had a much higher selling price and profit. This was consistent in the head office and product categories.

The results of the profit optimisation showed that employing six baristas would be optimal for both coffee shops 1 and 2. This took into account the number of customers that could be served versus the number of baristas employed and selected that number based on the capacity of the coffee shop and the point of maximum profit.

Anova

The objective was to find out if a difference appeared between the different years' average delivery hours for every product category. This would provide an accurate estimate of whether or not the delivery performance improved or declined from one year to the next. This information would be useful for the company when planning strategies for the following year.

This was done by calculating a one-way ANOVA for each product category. The null hypothesis was rejected if the probability was smaller than 0.05.

H_0 = identical mean delivery times between the two years

H_1 = the difference present between the two years

Product	**P-value**	**Description**
Cloud subscription	0.86	Null hypothesis is not rejected and there is not a significant difference in delivery times between the 2 years
Software	0.67	Null hypothesis is not rejected and there is not a significant difference in delivery times between the 2 years
Keyboard	0.00451	Null hypothesis is rejected and there is a significant difference in delivery times between the 2 years
Mouse	0.467	Null hypothesis is not rejected and there is not a significant difference in delivery times between the 2 years
Monitor	0.504	Null hypothesis is not rejected and there is not a significant difference in delivery times between the 2 years
Laptop	0.481	Null hypothesis is not rejected and there is not a significant difference in delivery times between the 2 years

Only the product keyboard had a statistical difference present. This means that most of the products had a stable average delivery. The delivery efficiency in the keyboard category could have been affected by various factors, such as supplier delays or process adjustments.

This is further supported by the mean standard deviation plots that can be seen in Figures 59-64. The average delivery hours are stable for most product categories. However, this is not typical of the keyboard category. In most categories, the mean delivery time was almost identical. This supports the ANOVA results.

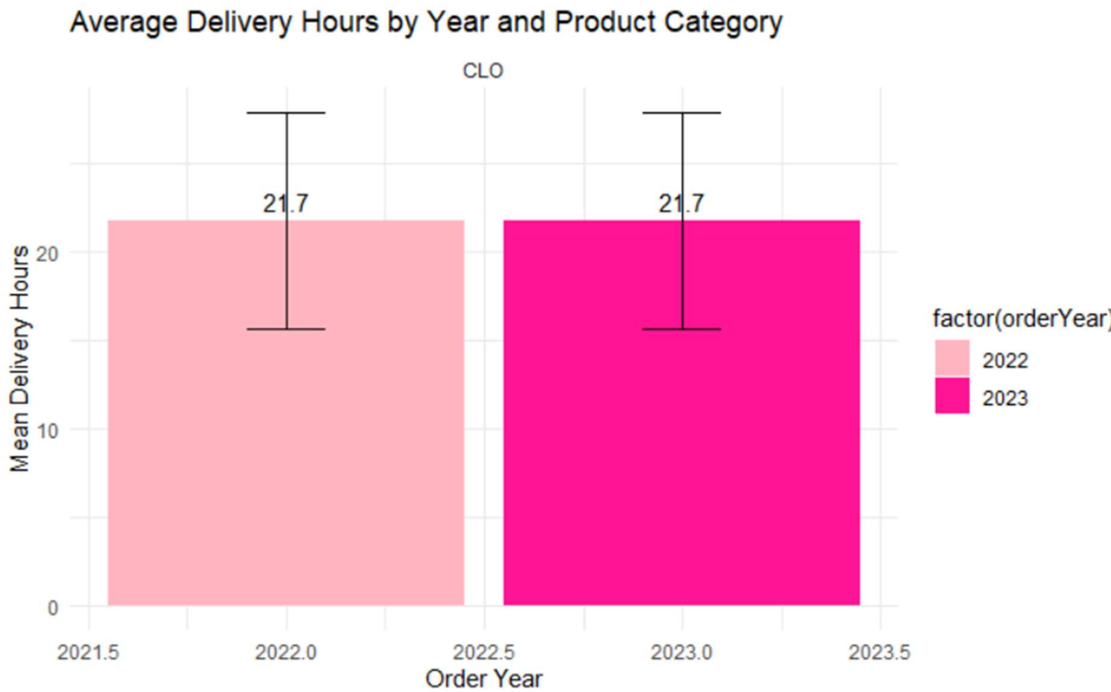

Figure 59: Mean standard deviation plot for cloud subscription

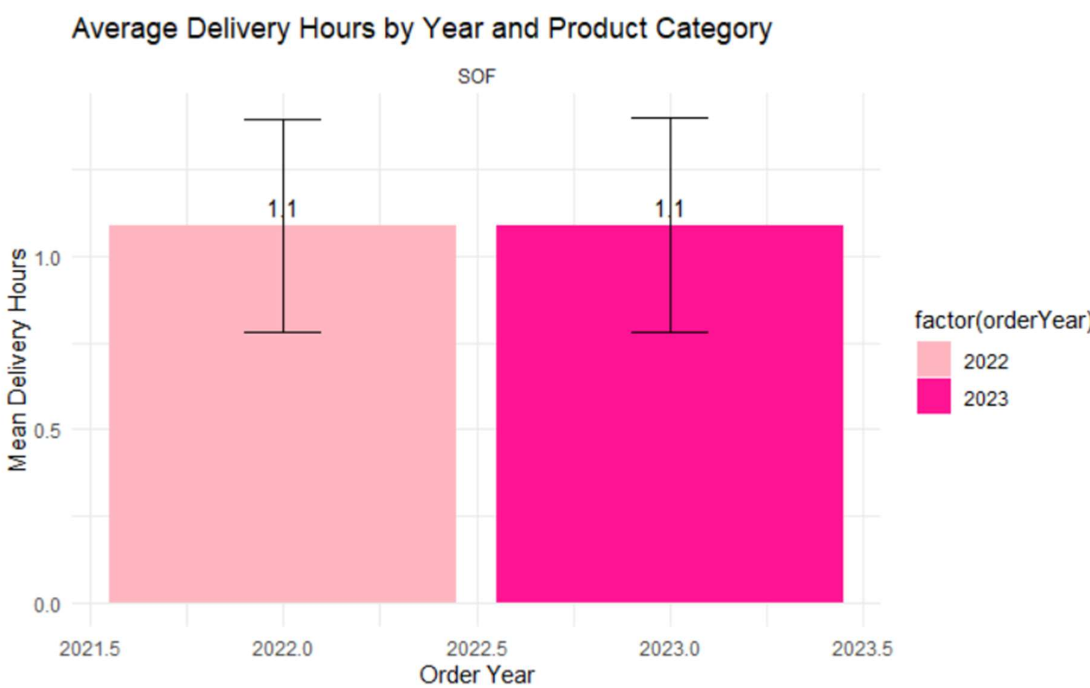

Figure 60: Mean standard deviation plot for software

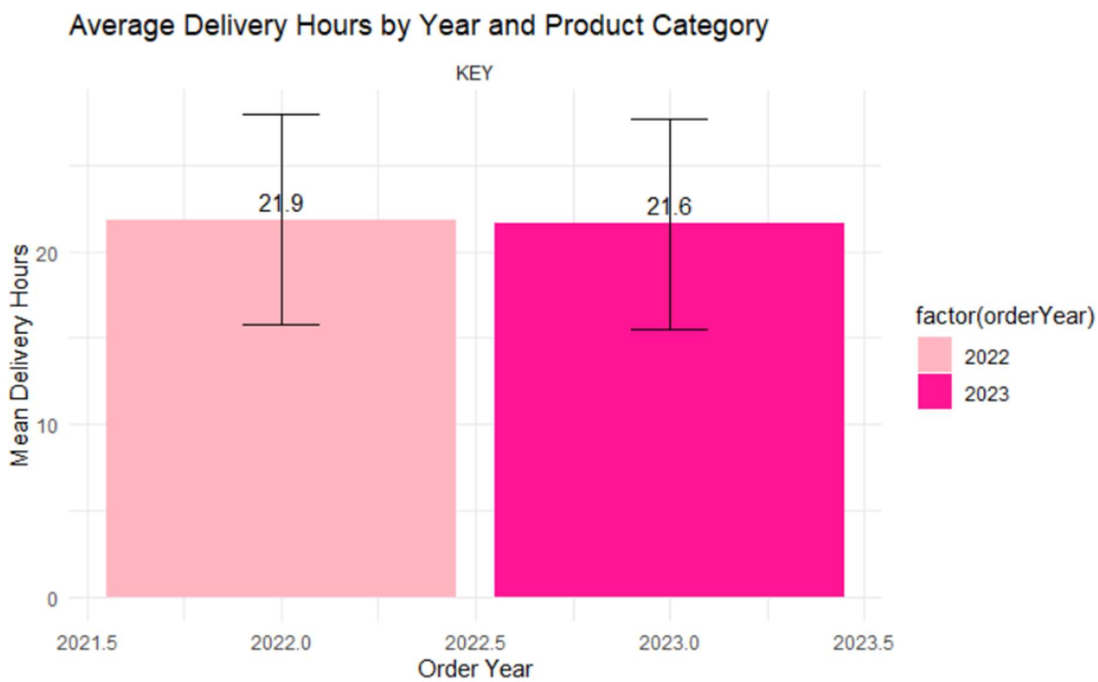

Figure 61: Mean standard deviation plot for keyboard

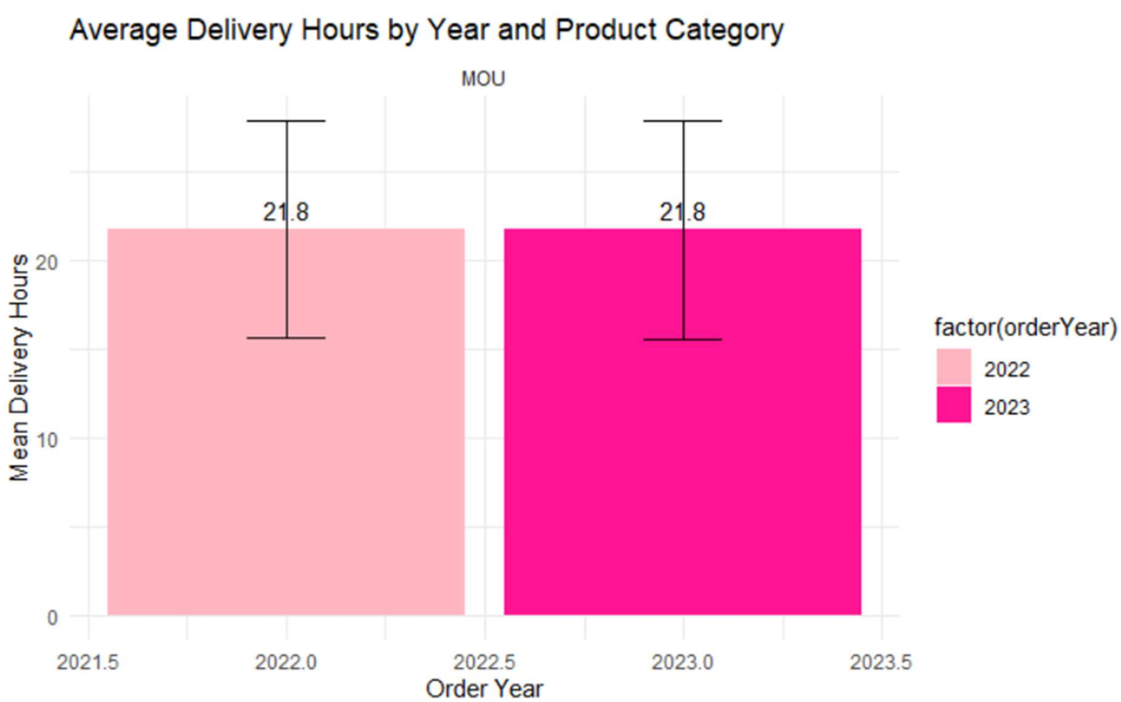

Figure 62: Mean standard deviation plot for mouse

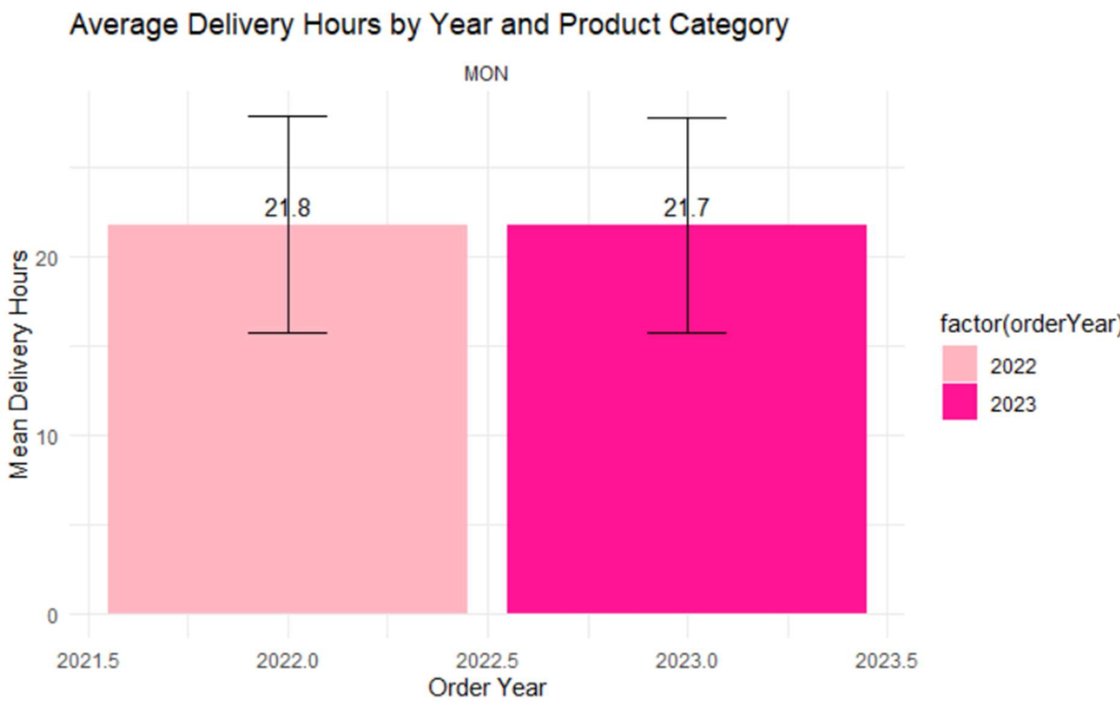

Figure 63: Mean standard deviation plot for monitor

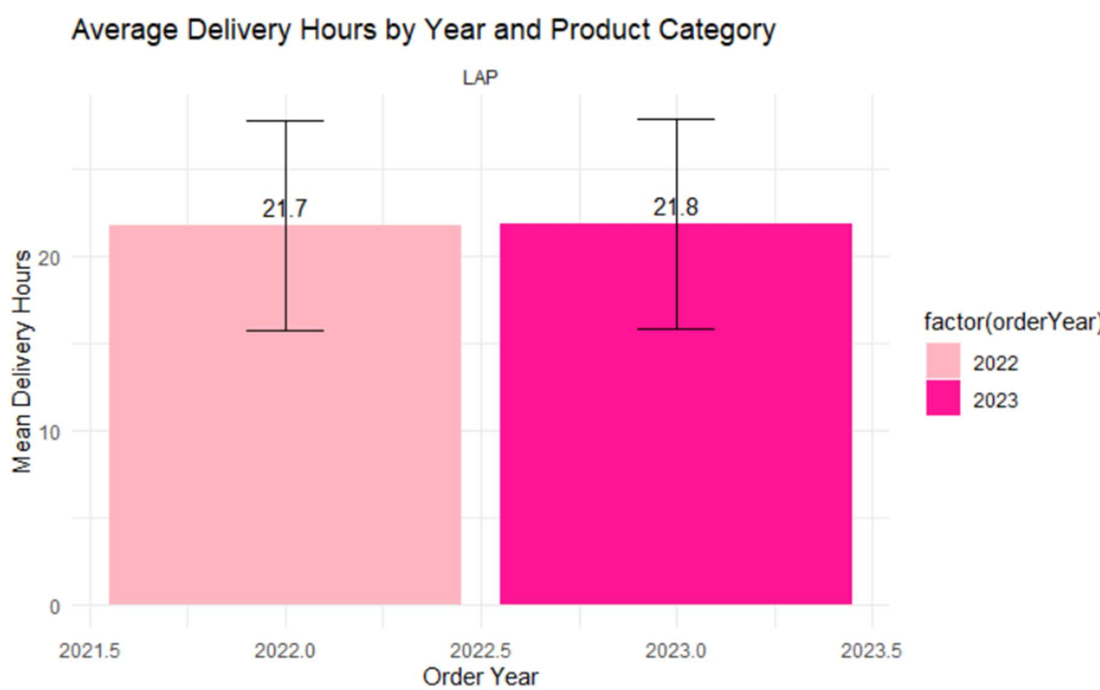

Figure 64: Mean standard deviation plot for laptops

Service Reliability

Reliable days

From the graph that was provided in the ESCA document, it was assumed that the car rental agency employed 16 people. The probability of one of these employees coming to work was calculated by multiplying the total number of employees summed by the total number of work days (397) and dividing this by the multiplying 16 by 397. This resulted in a probability of 0.974. This would indicate that an employee would work 9.49 days out of a 365 day year. The business was considered unreliable if there were less than 15 employees working on any given day. Using the binomial formula, the percentage of reliable days was 0.936 which would mean that 341.8 days of the year were reliable.

Profit optimisation

The optimal number of employees that the company should employ was determined using a cost model that aided minimising the total annual costs and taking the expenses of sale losses and employee costs into account.

Each employee is paid R25 000 a month. While employing more people increases the cots of the company, it also means that the risk of service failure is avoided.

In Figure 61, it is clear that there is a linear increase from R0 to R300000, as the number of employees increase from 16 to 17. IN this period, there is also a significant decreased in the loss of sales. The total cost also reaches its lowest point at 17 employees. This would strongly suggest that 17 employees would enable the maximum profit for the company.

Employing more than 17 people increases the total cost beyond the benefit of the smaller loss in sales as the loss in sales begins to plateau. This means that beyond 17 employees, the capacity of the company has already been reached so there are no extra sales coming in to reduce the total costs.

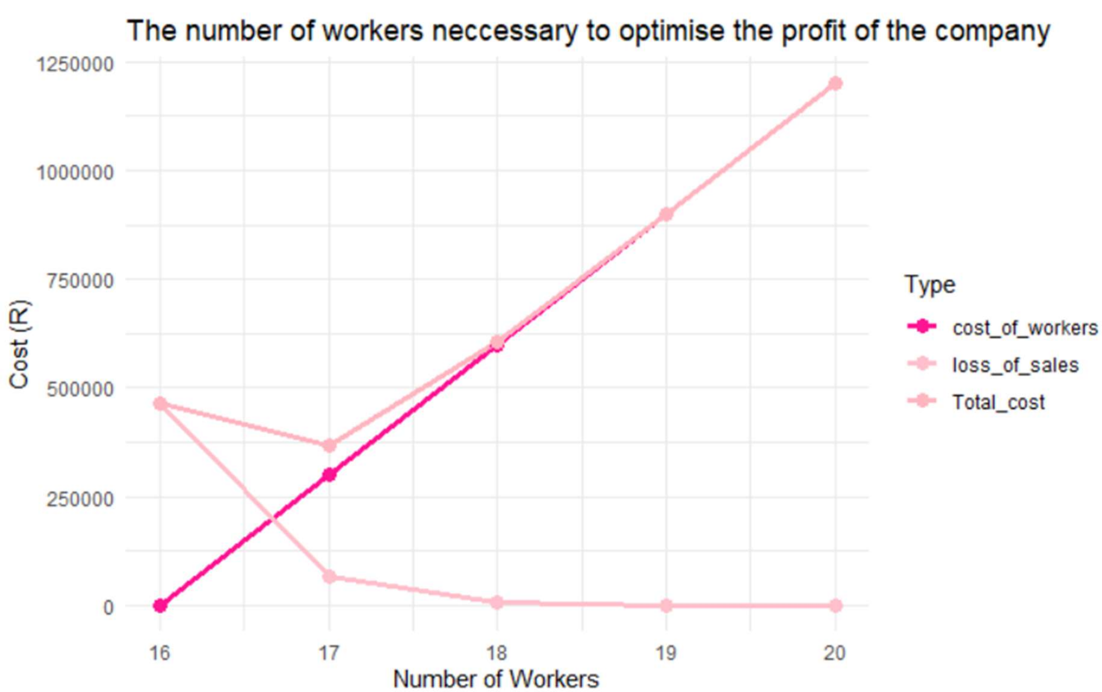

Figure 65: profit optimisation through the number of employees

Conclusion

The data analysis that was undertaken provided the company with an in-depth overview of opportunities for improvement and operational performance. The analysis of customers showed that the overall performance is influenced more greatly by the income of the customers rather than by their age. Further, middle-aged classes have a much stronger purchasing power as they have larger incomes than the other age groups.

Investigating the sale of products shows that selling prices and markups are much more standardised internally at the head office than they are seen to be in the market that experiences seasonal fluctuations. This gives the company an opportunity to do strategic planning that considers holiday seasons to increase the profit it makes. After data correction, it was evident that product categories such as monitors and laptops are generally more expensive and bring in the largest profit for the company.

The investigation of the process capabilities revealed that most products remained within the control limits, but consistent small deviations could be caused by systematic issues within the company. The statistical process control analysis showed that the delivery times are gradually increasing, more specifically for physical products. This was supported by the ANOVA results.

Service reliability studies assisted the car rental agency and the coffee shops in identifying the optimal number of employees to employ, based on the capacity of the companies, the point at which the most profit was made, and at what point too many employees would begin to negatively impact the overall profit.

Overall, data analysis should be prioritised by all companies to find and address process control issues, find patterns in sales that could help marketing strategies, and best calculate how to maximise profits.

References

1. OpenAI. 2025. *ChatGPT (GPT-5)*, Version October 2025 [Software]. Available: https://chat.openai.com/ [2025, October 23].
2. Dirkse van Schalkwyk, T. 2025. *Short summary of SPC (Statistical Process Control) and Limits.* Stellenbosch University [Online]. Available: https://learn.stem.ac.za/ [2025, October 23].