Luke McMeeking – 23773537
Stellenbosch University
Department of Industrial Engineering
October 2022

ECSA Project 2022

Quality Assurance – 344

Table of Contents

List of Tables

List of Figures

Introduction

As part of the Quality Assurance 344 module in the Industrial Engineering course as Stellenbosch University, it was required that a report be submitted to the Engineering Counsel of South Africa that is able to prove graduate attribute 4. Given client data from an online business, the goal of this report is to provide analysis on the data, using statistical analysis and other quality control measures to provide necessary information, recommendations and conclusions to the stakeholders and management of the business.

To accomplish this goal, as set out in the report, a number of steps were taken. These included data wrangling, in which the data was cleaned by removing missing and invalid instances, statistical analysis to become more familiar with the data and identify potential patterns and trends, statistical process control to measure quality control, optimising the delivery process, finding relationships in the data using DOE and MANOVA, and measuring the reliability of the service and products.

Data Wrangling

The original dataset had 10 descriptive features and consisted of 180 000 instances. In order to correctly analyse the given client data, it was essential that the data was first cleaned so that only valid data was used later on. Using data wrangling, a number of invalid instances were identified and removed from our data set. After wrangling with the data, 22 invalid instances were found. Of these 22 invalid instances, 17 of them were NA price values and the remaining 5 were negative price values. These instances are listed in Table 1 below.

Table 1 – Invalid Data

NUM	X	ID	AGE	Class	Price	Year	Month	Day	Delivery.time	Why.Bought
1	98765	64288	25	Clothing	NA	2021	1	24	8.5	Browsing
2	54321	62209	34	Clothing	NA	2021	3	24	9.5	Recommended
3	34567	18748	48	Clothing	NA	2021	4	9	8.0	Recommended
4	155555	33583	56	Gifts	NA	2022	12	9	10.0	Recommended
5	144443	37737	81	Food	-588.8	2022	12	10	2.5	Recommended
6	177777	68698	30	Food	NA	2023	8	14	2.5	Recommended
7	16320	44142	82	Household	-588.8	2023	10	2	48.0	EMail
8	56789	63849	51	Gifts	NA	2024	5	3	10.5	Website
9	19998	68743	45	Household	-588.8	2024	7	16	45.5	Recommended
10	87654	40983	33	Food	NA	2024	8	27	2.0	Recommended
11	166666	60188	37	Technology	NA	2024	10	9	21.5	Website
12	19541	71169	42	Technology	NA	2025	1	19	20.5	Recommended
13	19999	67228	89	Gifts	NA	2026	2	4	15.0	Recommended
14	155554	36599	29	Luxury	-588.8	2026	4	14	3.5	Recommended
15	12345	18973	93	Gifts	NA	2026	6	11	15.5	Website
16	23456	88622	71	Food	NA	2027	4	18	2.5	Random
17	65432	51904	31	Gifts	NA	2027	7	24	14.5	Recommended
18	144444	70761	70	Food	NA	2027	9	28	2.5	Recommended
19	19540	65689	96	Sweets	-588.8	2028	4	7	3.0	Random
20	76543	79732	71	Food	NA	2028	9	24	2.5	Recommended
21	16321	81959	43	Technology	NA	2029	9	6	22.0	Recommended
22	45678	89095	65	Sweets	NA	2029	11	6	2.0	Recommended

Thus, the remaining 179978 instances were valid and were used for the further analysis. Table 2 below contains a summary of all the data before removing any invalid instances.

Table 2 – Summary of All Sales Data

X	ID	AGE	Class	Price	Year
Min. : 1	Min. :11126	Min. : 18.00	Length:180000	Min. : -588.8	Min. :2021
1st Qu.: 45001	1st Qu.:32700	1st Qu.: 38.00	Class :character	1st Qu.: 482.3	1st Qu.:2022
Median : 90000	Median :55081	Median : 53.00	Mode :character	Median : 2259.6	Median :2025
Mean : 90000	Mean :55235	Mean : 54.57		Mean : 12293.7	Mean :2025
3rd Qu.:135000	3rd Qu.:77637	3rd Qu.: 70.00		3rd Qu.: 15270.7	3rd Qu.:2027
Max. :180000	Max. :99992	Max. :108.00		Max. :116619.0	Max. :2029
				NA's :17	

Month	Day	Delivery.time	Why.Bought
Min. : 1.000	Min. : 1.00	Min. : 0.5	Length:180000
1st Qu.: 4.000	1st Qu.: 8.00	1st Qu.: 3.0	Class :character
Median : 7.000	Median :16.00	Median :10.0	Mode :character
Mean : 6.521	Mean :15.54	Mean :14.5	
3rd Qu.:10.000	3rd Qu.:23.00	3rd Qu.:18.5	
Max. :12.000	Max. :30.00	Max. :75.0	

Table 3 below, shows the summary of only the valid data, after all invalid instances were removed. An extra column called NUM was also added to keep track of the new order of instances since 22 instances were removed. This is the data that was used in the rest of the report.

Table 3 – Summary of Valid Data

NUM	X	ID	AGE	Class	Price
Min. : 1	Min. : 1	Min. :11126	Min. : 18.00	Length:179978	Min. : 35.65
1st Qu.: 44995	1st Qu.: 45004	1st Qu.:32700	1st Qu.: 38.00	Class :character	1st Qu.: 482.31
Median : 89990	Median : 90004	Median :55081	Median : 53.00	Mode :character	Median : 2259.63
Mean : 89990	Mean : 90003	Mean :55235	Mean : 54.57		Mean : 12294.10
3rd Qu.:134984	3rd Qu.:135000	3rd Qu.:77637	3rd Qu.: 70.00		3rd Qu.: 15270.97
Max. :179978	Max. :180000	Max. :99992	Max. :108.00		Max. :116618.97

Year	Month	Day	Delivery.time	Why.Bought
Min. :2021	Min. : 1.000	Min. : 1.00	Min. : 0.5	Length:179978
1st Qu.:2022	1st Qu.: 4.000	1st Qu.: 8.00	1st Qu.: 3.0	Class :character
Median :2025	Median : 7.000	Median :16.00	Median :10.0	Mode :character
Mean :2025	Mean : 6.521	Mean :15.54	Mean :14.5	
3rd Qu.:2027	3rd Qu.:10.000	3rd Qu.:23.00	3rd Qu.:18.5	
Max. :2029	Max. :12.000	Max. :30.00	Max. :75.0	

Standard Descriptive Statistics

Descriptive Statistics

Descriptive statistics is used widely to help summarise and describe features or characteristics of a data set. When applied correctly, descriptive statistics can be a very useful business and decision-making tool. Before attempting to use descriptive statistics on the client data, it is important to know what each feature is and to understand the values it can take (excluding NUM, X and ID)

- AGE – age of client
  Ranges from 18 to 108 years old
- Class – type of product sold
  7 classes: Clothing, Household, Food, Technology, Sweets, Gifts, Luxury
- Price – how much is paid for a product
  Ranges from R35.65 to R116 618.97
- Year – what year the product was bought in
  Ranges from 2021 to 2029
- Month – what month the product was bought in
  Ranges from 1 to 12
- Day – what day the product was bought in
  Ranges from 1 to 30
- Delivery.time – how long it took for the product to be delivered
  Ranges from 0.5 to 75

- Why.bought – client's reason for buying the product
  6 classes: Browsing, Email, Random, Recommended, Spam, Website

In figure 1 below for the class distribution, it can be seen that classes the contribute most to the items sold is gifts and technology classes.

Figure 1 – Class Distribution

In the why bought distribution below (Figure 2), recommended is the reason that makes up the majority of the businesses sales. This is positive, as it means that customers are satisfied with the products they are receiving and are recommending it to others. However, the business could possibly improve other areas such as sending out more emails, improving their website, and improving the quality and frequency of advertisements so that more people see them when browsing.

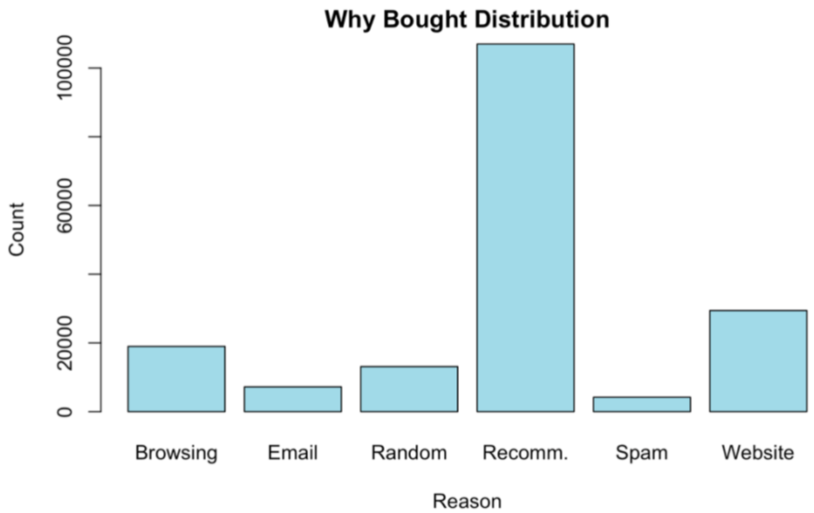


Figure 2 – Why Bought Distribution

Figure 3 further shows how that for all classes of items, recommended is the most frequent reason for why customers buy products from the company. However, for household, it is clear that the ratio between recommended and website is much closer than in any of the other classes.

Figure 3 – Histogram of Reason Bought per Class

The age distribution shown below (Figure 4), follows a skewed right distribution, meaning younger people are making more frequent purchases than older people. The age feature also goes up to 109 years old, which considering that there are not many people at all who live to be that old, this could represent some sort of error in the data, and should be investigated by the business.

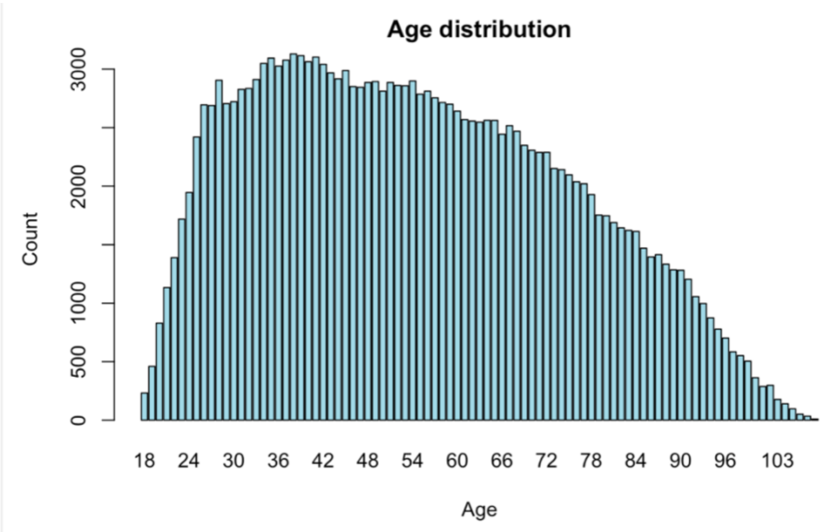


Figure 4 – Age Distribution

From the bar plot below of the average price of items per class (Figure 5), luxury items are, on average, the most expensive items to sold, followed by technology items. Sweets, food and clothing items are the least expensive. It is expected that luxury and technology items would be the most expensive.

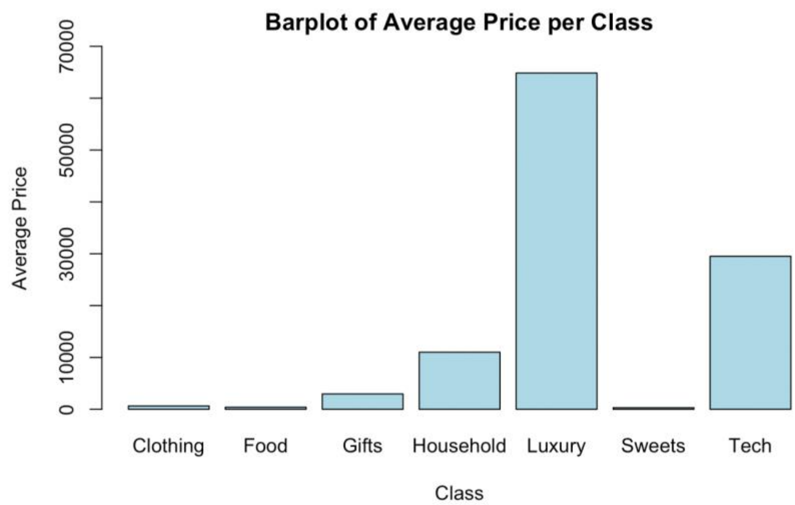


Figure 5 – Bar plot of Average Price per Class

As seen in the bar plot below (Figure 6), 2021 provided the largest amount of sales. In 2022, however, sales took a large dip and remained relatively constant up until 2027, when sales started following an upward trend. This dip in sales could possibly be due to decreased demand or external socio-economic factors. This needs to be investigated, and innovative methods should be implemented in order to boost sales again.

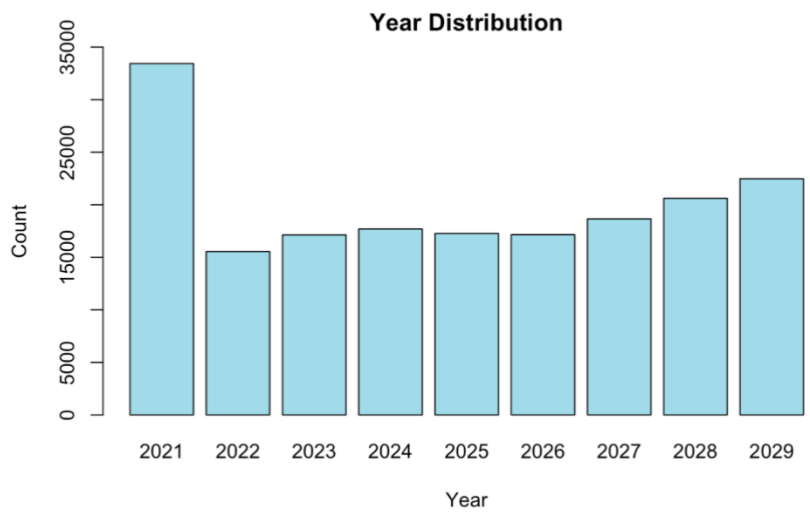


Figure 6 – Year Distribution

Household items, on average, take the longest out of all the different classes to be delivered (As seen in figure 7). Sweets, food and luxury all take the shortest. The delivery time for the household items should be investigated to see why this is the case and how the delivery time can be improved, as it is unclear as to why this is.

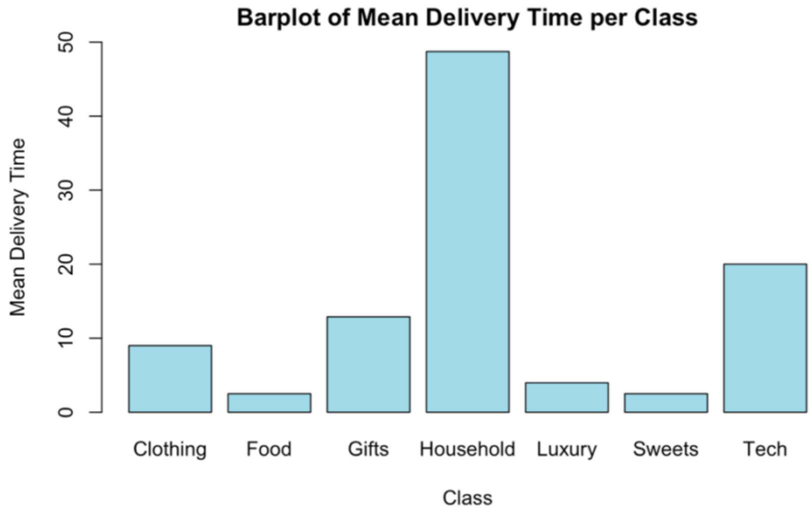


Figure 7 – Bar plot of Mean Delivery Time per Class

Figure 8 depicts histograms of delivery times for each class. From the histograms, it can be seen that the delivery times of all the classes follow a normal distribution. This relates to figure # that showed the average delivery time per class, and shows that household not only has the longest delivery times, but also the most varied, as seen in the lower count for each delivery time. On the other hand, food has one of the lowest delivery times, but least varied.

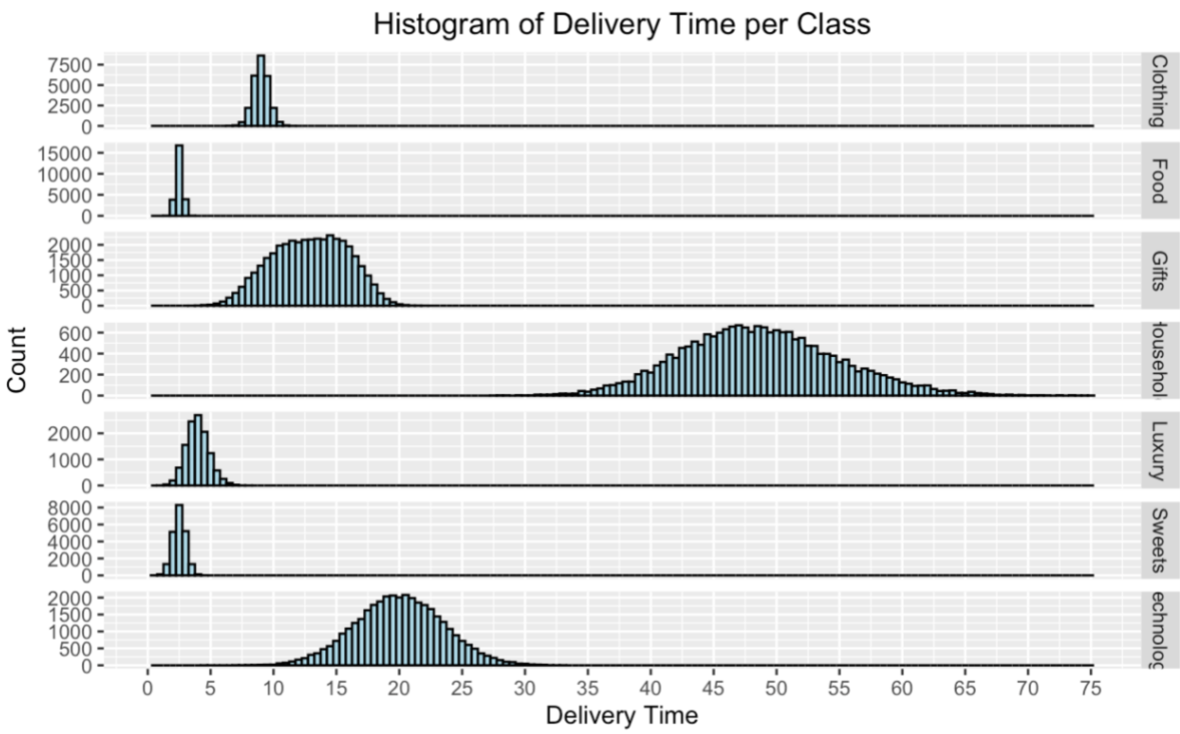


Figure 8 – Histogram of Delivery Time per Class

Figure 9 shows histograms of the price and its frequency for each class of item. As seen, clothing, food and sweets items have the lowest spread of price values, with the prices also being very low. Luxury items vary the most in terms of cost, and follow a uniform distribution. Household and technology also have a relatively large price spread, and also follow a uniform distribution.

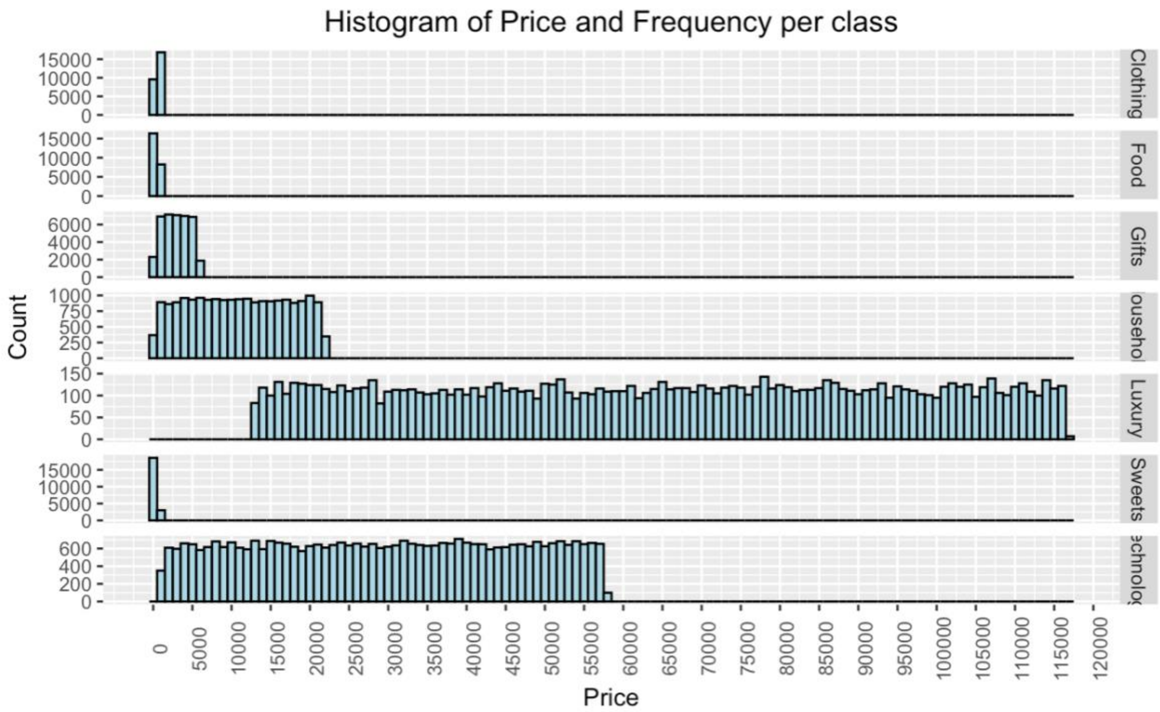


Figure 9 – Histogram of Price and Frequency per Class

Capability Indices

Process capability is the ability of a process to meet specifications. Two important measures, capability potential (C_P) and capability performance (C_{PK}) help to illustrate this process's ability to meet these given specifications. The C_P ratio helps to determine how well the process spread (defined as 6 standard deviations of the process) fits into the given specification range, while the C_{PK} ratio does the same thing, but additionally it measures how close the process mean is to the target value of the given specification.

When calculating the process capability indices for the process delivery times of the technology class products, it was given that the upper service level (USL) was 24 hours and the lower service level (LSL) was 0 hours. The following formulas were used to calculate the different process capability indices.

$$\boldsymbol{C_P = \frac{USL - LSL}{6\sigma} = \frac{24 - 0}{6(3.502)} = 1.1422}$$

$$\boldsymbol{C_{PU} = \frac{USL - \bar{X}}{3\sigma} = \frac{24 - 21.095}{3(3.502)} = 0.3796933}$$

$$\boldsymbol{C_{PL} = \frac{\bar{X} - LSL}{3\sigma} = \frac{21.095 - 0}{3(3.502)} = 1.90472}$$

$$\boldsymbol{C_{PK} = min(C_{PU}, C_{PL}) = 0.3796933}$$

A LSL of 0 is logical, as it is not possible for delivery times to be negative. A C_{PK} of 1 is deemed barely capable, while anything above 1.3 is considered capable. Higher ratios are always better, indicating that the number of non-conforming delivery times is very low.

Since the C_{PK} ratio is too low for the process, it would be advised that investigation is required in order to increase it to an acceptable level. It can be improved by recentring the process on the target, or reducing the variation and spread of the process by looking at and redesigning the process to perform better to what is required.

Statistical Process Control

Statistical process control (SPC) is a statistical method used to monitor and control the quality of a process, in order to ensure that the process operates correctly and efficiently. It is useful for identifying causes of variation and when a process goes out of control, and what can be done to correct it. SPC charts are a popular tool to record data and identify unexpected instances (variation).

The variation within a process, can be classified as one of two types. Common-cause variation is variation that is intrinsic to the process, characterised by stable, predictable and consistent variation over time, within specified limits. This variation is normal, and is an indication that a process is stable and in control. Special-cause variation indicates that a process is out of control, and is characterised by unpredictable variation outside of specified limits.

Initialising charts

In order to be able to calculate the initial control limits, the first 30 samples of 15 sales each were used to initialise the SPC charts and calculate the centre line, the upper control limits, the lower control limits, the 2-sigma control limits and the 1-sigma control limits. Figure 10 below shows an example of the initialisation of the $\bar{X}$-and-S SPC charts, namely for the clothing class.

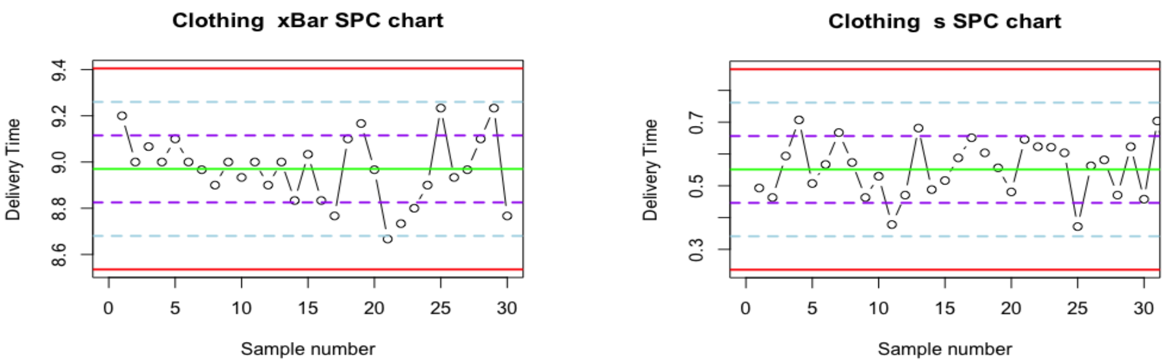


Figure 10 – Initialisation charts for Clothing

All the lines and limits calculated above, are listed in Table 4 for $\bar{X}$ and in Table 5 for S, for each class.

Table 4 - $\bar{X}$ Control Limit Table

	UCL	U2Sigma	U1Sigma	CL	L1Sigma	L2Sigma	LCL
Clothing	9.405	9.260	9.115	8.970	8.825	8.680	8.535
Household	50.248	49.019	47.791	46.562	45.333	44.105	42.876
Food	2.709	2.636	2.563	2.490	2.417	2.344	2.271
Technology	22.974	22.107	21.241	20.374	19.507	18.641	17.774
Sweets	2.897	2.757	2.618	2.478	2.338	2.199	2.059
Gifts	9.488	9.112	8.737	8.361	7.985	7.610	7.234
Luxury	5.494	5.241	4.989	4.736	4.483	4.231	3.978

Table 5 – S Control Limit Table

	UCL	U2Sigma	U1Sigma	CL	L1Sigma	L2Sigma	LCL
Clothing	0.867	0.762	0.656	0.551	0.446	0.341	0.236
Household	7.344	6.453	5.563	4.672	3.781	2.891	2.000
Food	0.437	0.384	0.331	0.278	0.225	0.172	0.119
Technology	5.181	4.553	3.924	3.296	2.667	2.039	1.410
Sweets	0.835	0.734	0.632	0.531	0.430	0.328	0.227
Gifts	2.246	1.974	1.701	1.429	1.157	0.884	0.612
Luxury	1.511	1.328	1.144	0.961	0.778	0.594	0.411

Control charts for all the data

After initialising the SPC charts to calculate the centre line, upper and lower control limits and the different sigma control limits, the rest of the client data can now be plotted to identify any potential special-cause variation and to understand whether the delivery times for each class are in or out of control.

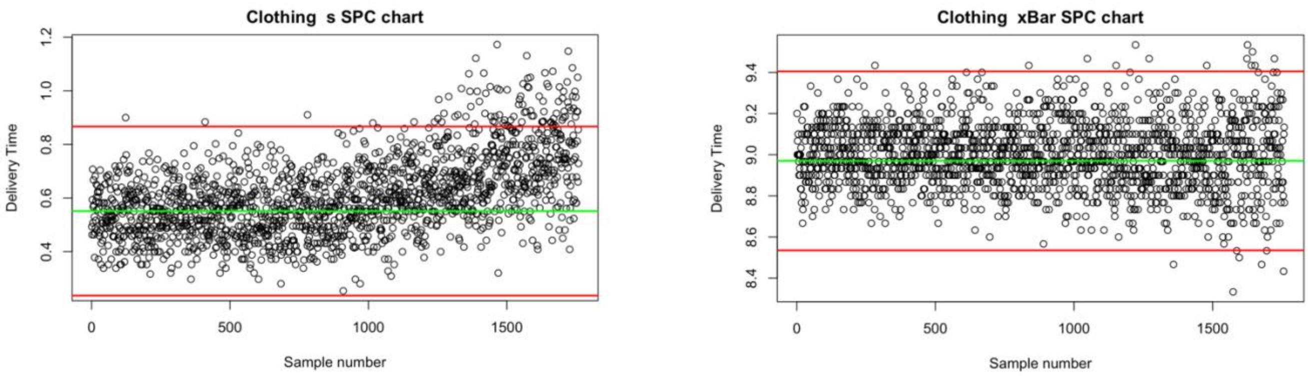


Figure 11 – Clothing S and $\bar{X}$ Control Charts

In figure 11 above for the clothing class, it can be seen that the standard deviation remained relatively in control for the majority of the samples, with only a few instances being above or below the outer control limits. However, in figure 11 above, after the first 1000 instances, the sample mean followed an upward trend, going above the upper control limit. So despite the distribution for the clothing delivery time remaining fairly constant for the majority of the samples, it exhibits signs of increasing mean delivery time.

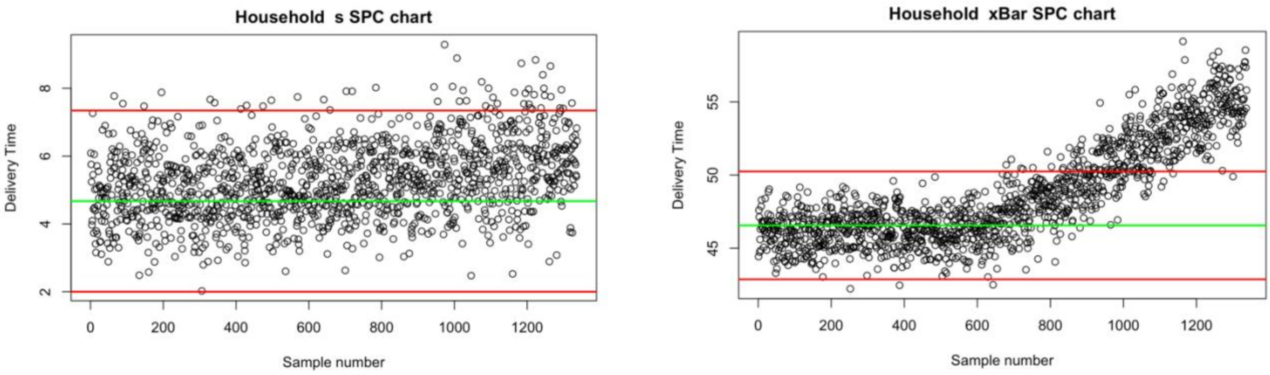
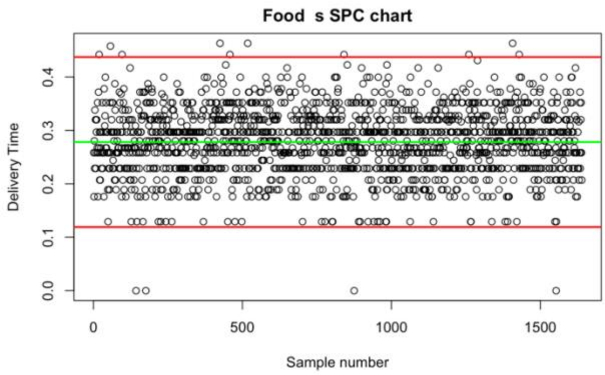


Figure 12 – Household S and $\bar{X}$ Control Charts

In figure 12 above for the household class, the standard deviation seems to be fairly in control, however there are more samples on one side of the centre line than the other. In figure 12, after around sample 800, the sample mean followed an upward trend, going very out of control above the upper control limit. Up until this point, the process seemed to be fine and in control.



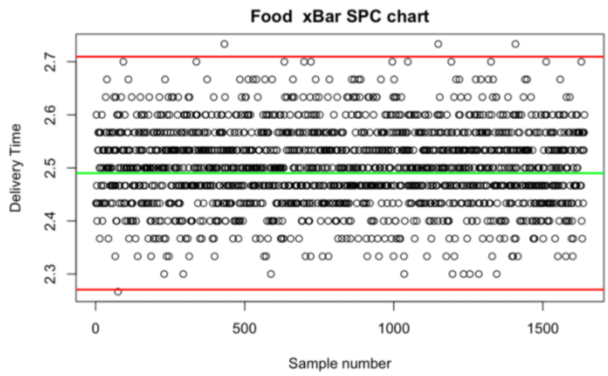
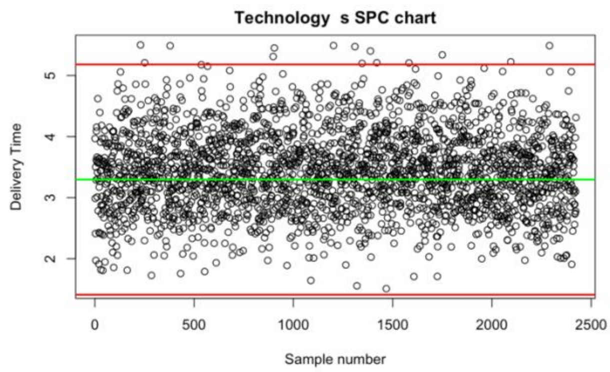


Figure 13 – Food S and $\bar{X}$ Control Charts

In figure 13 above for the food class, the standard deviation is in control, barring a few instances outside of the outer control limits. The sample mean in figure 13 above, stayed in control with only a few instances outside of the outer control limits.



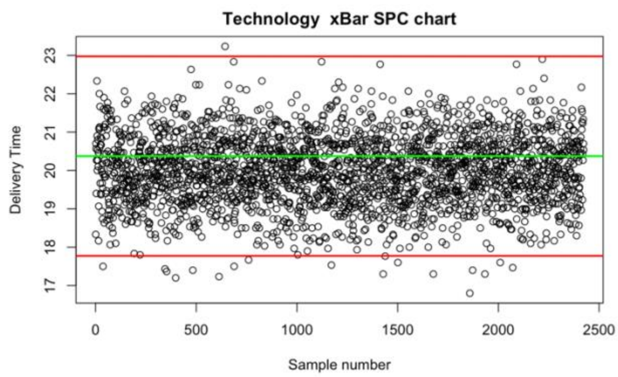
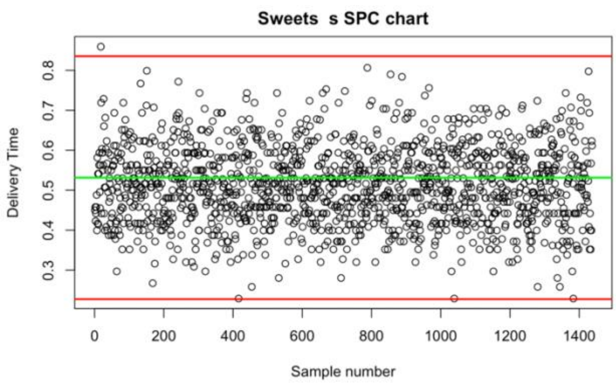


Figure 14 – Technology S and $\bar{X}$ Control Charts

In figure 14 above, the standard deviation for the technology class was in control, except for a few instances outside of the upper control limit. The sample mean in figure 14 stayed in control, with the spread of instances within the control limits remaining fairly constant, despite a few instances outside of the control limits.



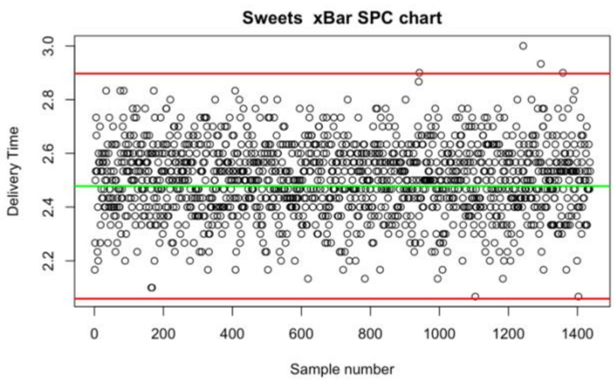


Figure 15– Sweets S and $\bar{X}$ Control Charts

In figure 15 above, both the standard deviation and sample mean remained in control for the delivery times for the sweets class. Both have a fairly even spread, with only a few instances outside of the control limits.

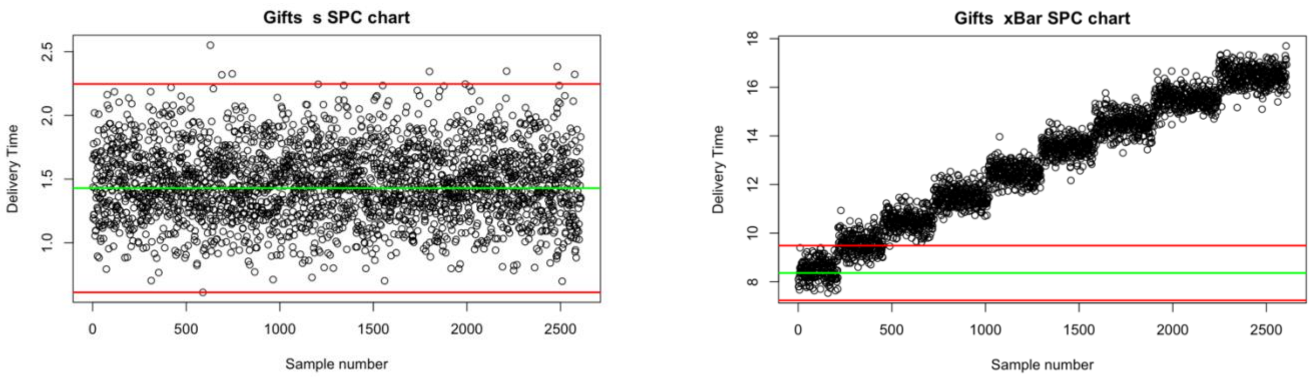


Figure 16 – Gifts S and $\bar{X}$ Control Charts

The standard deviation for the gifts class (figure 16 above) remained in control, however, the sample mean (also figure 16 above) went very out of control above the upper control limit, following a constant upward trajectory. This would be something that requires urgent attention.

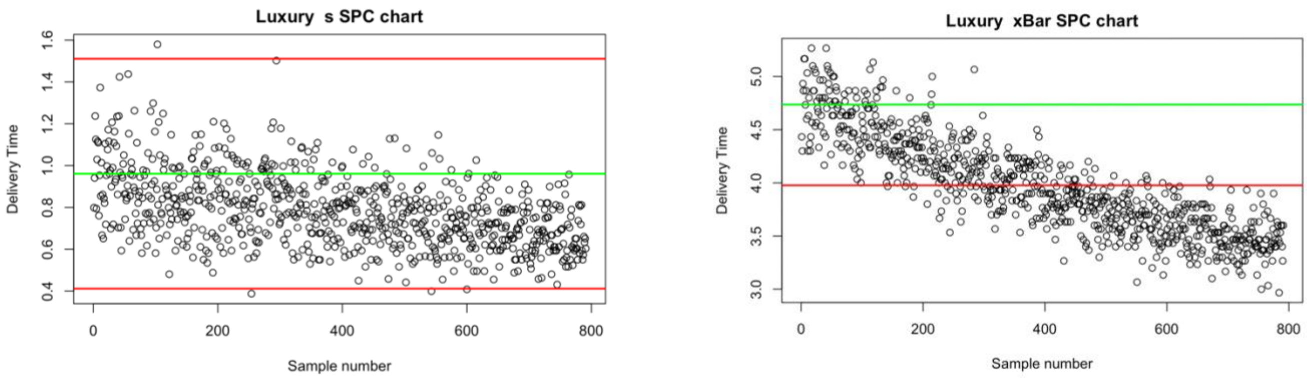


Figure 17 – Luxury S and $\bar{X}$ Control Charts

As seen in figure 17 above, many instances for the standard deviation for the luxury class fall below the centre line. In figure 17 above, it is clear to see that the process follows a downward trend and goes out of control.

Optimising the Delivery Process

$\overline{X}$ outside of Upper control limits

A number of instances were identified for each class that were outside of the outer control limits for the sample means. The total number of instances per class are listed in table 6 below. As seen in the table, and backed up by the graphs above, food and technology contained the least number of instances outside of the outer control limits (4 each), while gifts contained the most instances outside of the outer control limits (2287). Household and luxury also contained a large number of instances, 395 and 440 respectively, outside of the outer control limits.

Table 6 – Number of $\overline{X}$ instances outside of Outer Control Limits

	Class	Instances
1	Clothing	20
2	Household	395
3	Food	4
4	Technology	19
5	Sweets	4
6	Gifts	2287
7	Luxury	440

As well as the number of instances out of control, the first and last three instances are listed in table 7 below. Note that for the Food and Sweets classes, that only contain 4 instances each, the values listed are just the 4 instances.

Table 7 – First and Last 3 Instances of $\overline{X}$ outside Outer Control Limits

	Clothing	Household	Food	Technology	Sweets	Gifts	Luxury
1	282	252	75	37	942	213	142
2	837	387	432	345	1243	216	171
3	1048	643	1149	353	1294	218	184
4	1695	1335	1408	1933	1358	2607	789
5	1723	1336	NA	2009	NA	2608	790
6	1756	1337	NA	2071	NA	2609	791

As stated before, instances outside of the outer control limits can be an indication that special-cause variation is present and a process is out of control. It is important that once out of control samples are identified, that they are investigated to see whether outside factors are affecting the process and corrective action is required to make the process stable once more, or if there is nothing to worry about.

Sample Standard Deviations between -0.3 Sigma and +0.4 Sigma control limits
It was required that the most consecutive sample standard deviations between the -0.3 sigma and +0.4 sigma control limits were calculated, tabled and plotted. Table 8 below, lists the number of sample standard deviations identified between the -0.3 sigma and +0.4 sigma control limits for each class. As seen, the most consecutive instances were in the Gifts class, while the least were in the Household class.

Table 8 – Number of consecutive S Instances between -0.3 and +0.4 Sigma Control Limits

	Class	Instances
1	Clothing	4
2	Household	3
3	Food	5
4	Technology	6
5	Sweets	4
6	Gifts	7
7	Luxury	4

Table 9 below, shows the instance numbers for each class. Note that clothing, household, sweets and luxury consecutive samples occurred quite early on, while technology and gifts occurred closer to the end.

Table 9 – Consecutive S Instances

	1	2	3	4	5	6	7
Clothing	220	221	222	223	NA	NA	NA
Household	43	44	45	NA	NA	NA	NA
Food	752	753	754	755	756	NA	NA
Technology	1771	1772	1773	1774	1775	1776	NA
Sweets	91	92	93	94	NA	NA	NA
Gifts	2471	2472	2473	2474	2475	2476	2477
Luxury	60	61	62	63	NA	NA	NA

Below (Figure 18) are the plots for each class showing the consecutive instances between -0.3 sigma and +0.4 sigma for each class, as in table 9 above.

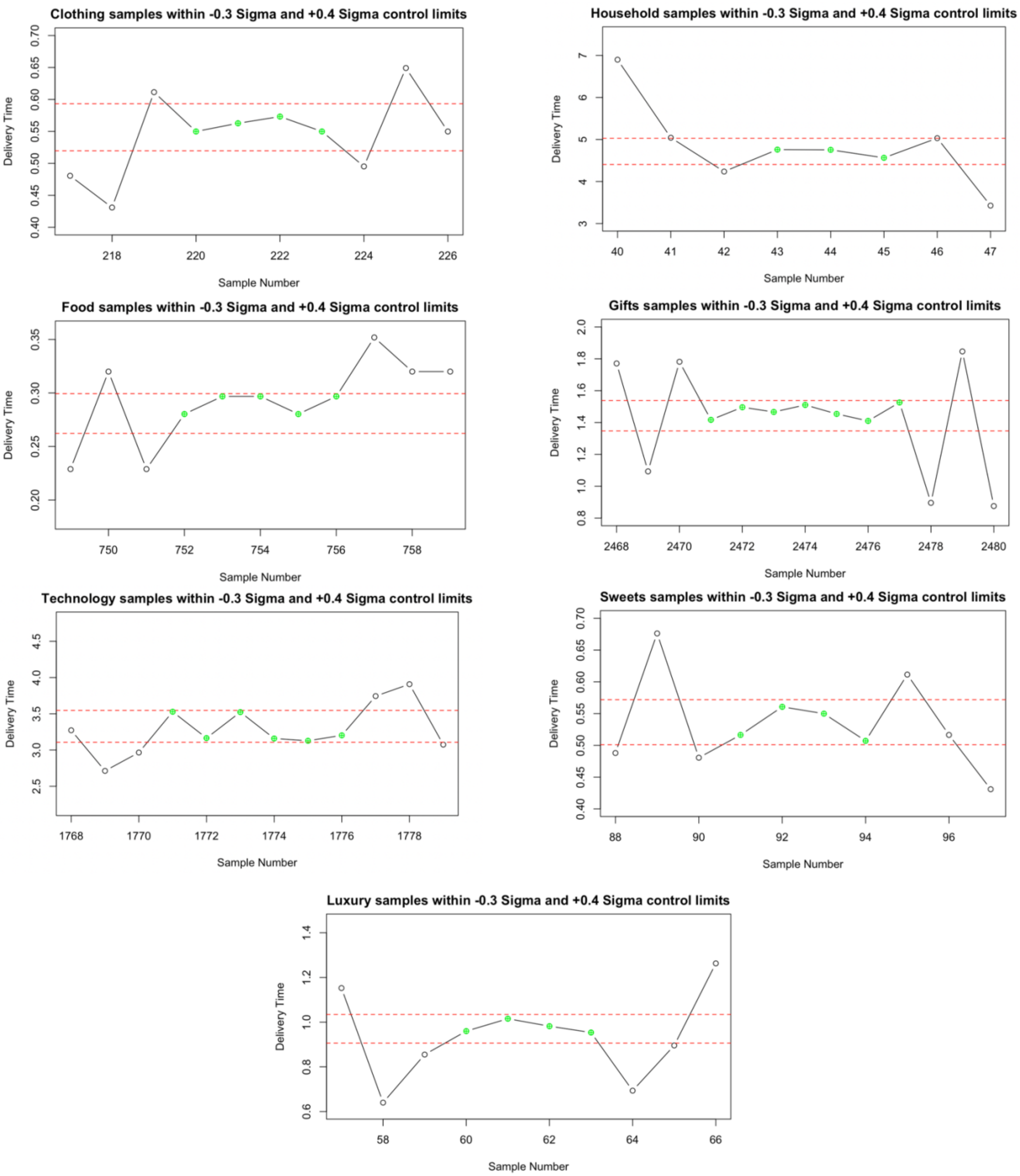


Figure 18 – Plots of all Consecutive S Instances between -0.3 Sigma and +0.4 Sigma for Each Class

Type 1 Error

A type 1 error, also known as producer's risk, manufacturer's risk or false alarm, is rejecting the null hypothesis when it's actually true.

In the case of the sample means outside of the outer control limits, the null hypothesis is that the process is in control (within the outer control limits), whereas the alternative hypothesis is that the process is not in control (outside the outer control limits). Thus it is clear to see, that a type 1 error would be saying that the process is out of control when it is actually in control.

Alpha = 0.002699796, meaning that there is a 0.27% chance that the process is flagged as out of control when it is actually in control.

In the case of the consecutive standard deviation samples between the -0.3 sigma and +0.4 sigma control limits, the null hypothesis is that the instances fall within the control limits, and the alternative hypothesis is that the instances fall outside of the control limits.

Optimising the delivery process for individual delivery times

It was given that for the technology class, products delivered slower than 26 hours incurred a R329/hour lost sales cost. In order to minimise the amount of lost sales revenue, it is possible to reduce the average time by an hour, with a cost of R2.50 per item per hour. In order to maximise profit, the delivery time should be shifted down by 3 hours. As can be seen in figure 19 below, shifting the delivery time down by 3 hours provides the least cost to the business.

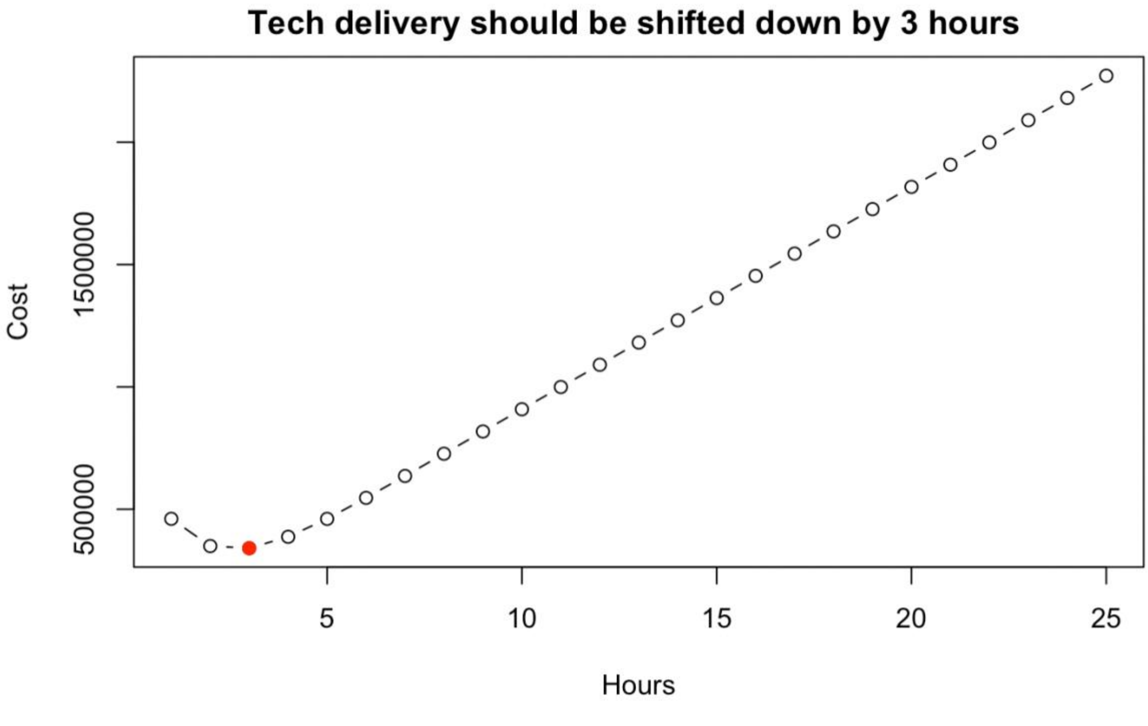


Figure 19 – Technology Delivery Optimisation Graph

Type 2 Error

A type 2 error, also known as consumers risk or misdetection, is accepting the null hypothesis when it is actually false. In the case of the type 2 error in the technology class, the null hypothesis would be that the sample mean value is between the UCL and LCL. Thus the type 2 error, would be saying that the sample mean value is between the UCL and LCL, when in reality it is actually outside of these outer control limits.

Since the mean has moved to 23 hours, the type 2 error is 0.4883. This is illustrated in the graph below, by the area within the control limits. Here the red lines indicate the lower control limit and the upper control limit respectfully. This means that if the process mean has moved to 23 hours, there is a 48.83% chance that the sample mean value will be out of control but incorrectly identified as in control.

Figure 20 – Type 2 Error

DOE & MANOVA

Multivariate Analysis of Variance (MANOVA) is used to determine the effect that groups of independent variables have on other dependant variables.

The first MANOVA table (Table 10) compares variables delivery time, year and price against class to see if there is any distinct relationship between them.

Table 10 – First MANOVA Table

```
 Response Delivery.time :
                Df   Sum Sq Mean Sq F value    Pr(>F)
Class            6 33458565 5576427  629429 < 2.2e-16 ***
Residuals   179971  1594452       9
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

 Response Year :
                Df  Sum Sq Mean Sq F value    Pr(>F)
Class            6  153040 25506.6  3698.2 < 2.2e-16 ***
Residuals   179971 1241263     6.9
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

 Response Price :
                Df     Sum Sq    Mean Sq F value    Pr(>F)
Class            6 5.7168e+13 9.5281e+12   80258 < 2.2e-16 ***
Residuals   179971 2.1366e+13 1.1872e+08
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

The second MANOVA table (Table 11) compares variables delivery time, year and price against the reason why bought to see if there is any distinct relationship between them.

Table 11 – Second MANOVA Table

```
 Response Delivery.time :
                Df   Sum Sq Mean Sq F value    Pr(>F)    
Why.Bought       5   783320  156664  822.74 < 2.2e-16 ***
Residuals   179972 34269697     190                      
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

 Response Year :
                Df  Sum Sq Mean Sq F value    Pr(>F)    
Why.Bought       5    3191  638.21  82.567 < 2.2e-16 ***
Residuals   179972 1391111    7.73                      
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

 Response Price :
                Df     Sum Sq    Mean Sq F value    Pr(>F)    
Why.Bought       5 1.5742e+12 3.1484e+11  736.26 < 2.2e-16 ***
Residuals   179972 7.6960e+13 4.2762e+08                      
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Reliability of the service and products

Problem 6 and 7

For Lafrideradora, one of the company's subsidiaries, it was given that the thickness of a refrigerator part is 0.06 ± 0.04 cm. Additionally, it costs \$45 to scrap a part outside of the given specifications.

$$L = k(y - m)^2$$
$$45 = k(0.04)^2$$
$$k = 28125$$

$$\Longrightarrow L(x) = 28125(y - 0.06)^2$$

The Taguchi loss function is plotted below (Figure 21). It is clear from the graph, that the closer the function is to 0.06, the lower the cost will be.

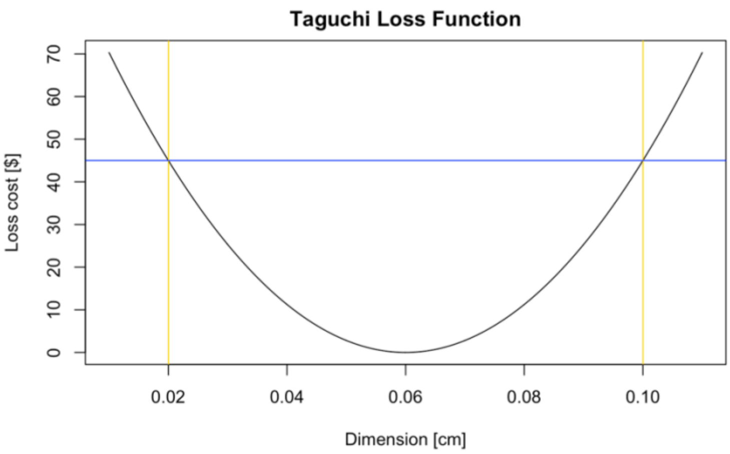


Figure 21 – Taguchi Loss Graph 1

After the scrap cost was reduced to 35$, the Taguchi loss function was calculated as follows, and plotted in figure 22.

$L = k(y - m)^2$
$35 = k(0.04)^2$
$k = 21875$

$$\Longrightarrow L(x) = 21875(y - 0.06)^2$$

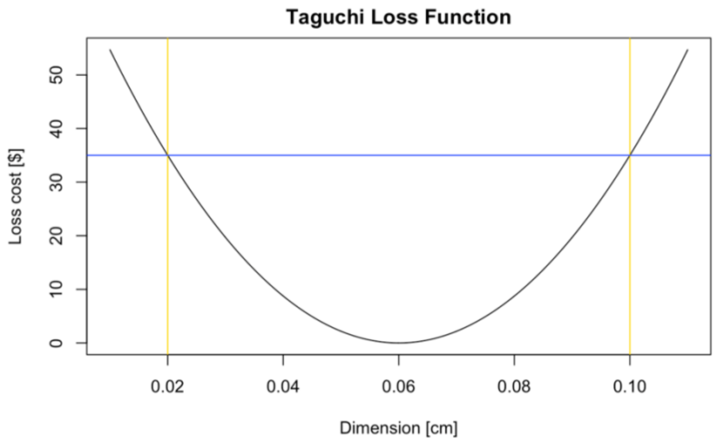


Figure 22 – Taguchi Loss Graph 2

If the process deviation from the target is reduced to 0.027cm, then the Taguchi Loss is calculated as follows, and plotted in Figure 23 below.

$L = 21875(0.027)^2 = \$15.95$

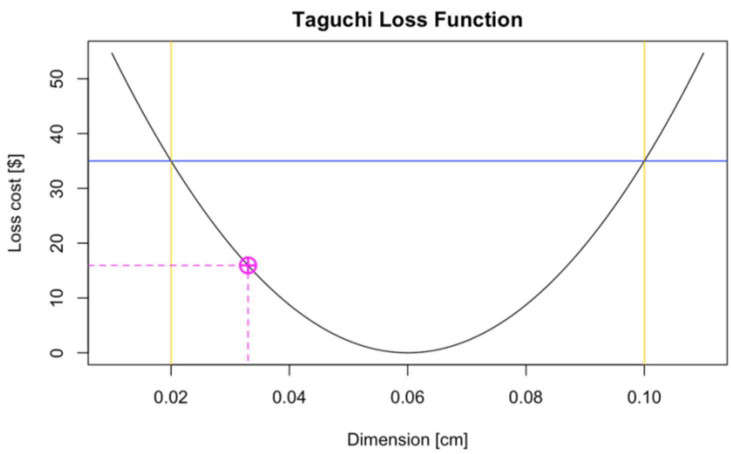


Figure 23 – Taguchi Loss Graph 3

Problem 27
Magnaplex, another subsidiary of the company, is in charge of manufacturing some of the company's technology products. On the orders of management, their current process needs to be investigated to see if it is wasteful due to the identical backup machines used in the event of failure.

Their current set-up looks as follows (Figure 24).

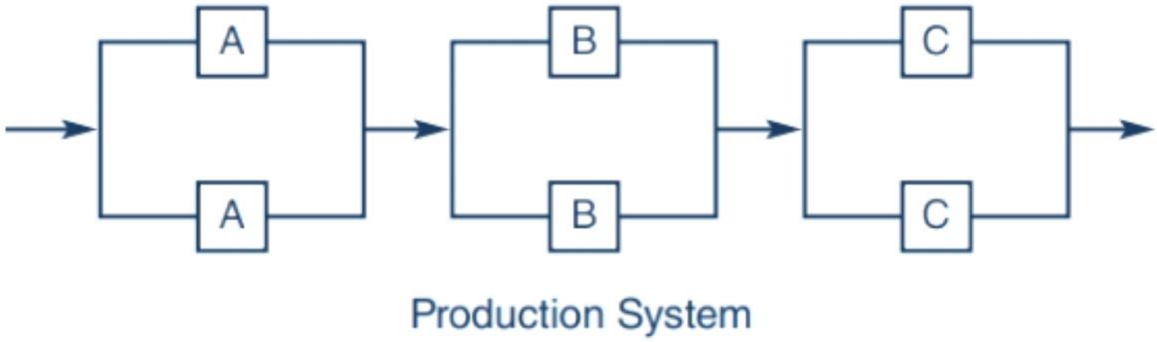


Figure 24 – Magnaplex Production System

The reliabilities of A, B and C is 0.85, 0.92 and 0.90 respectively. Initially they only have one of each machine (A, B and C) running at one time, and the system reliability is calculated as follows.

$$System\ Reliability = R_A R_B R_C$$
$$System\ Reliability = (0.85)(0.92)(0.90) = 0.7038 = 70.38\%$$

In order to increase the system reliability, all of the machines are now running at once, and the system reliability is calculated as follows.

$$P(Both\ A's\ fail) = (1 - 0.85)^2 = 0.0225$$
$$P(Both\ B's\ fail) = (1 - 0.92)^2 = 0.0064$$
$$P(Both\ C's\ fail) = (1 - 0.90)^2 = 0.01$$
$$Combined\ reliability_A = 1 - 0.0225 = 0.9775$$
$$Combined\ reliability_B = 1 - 0.0064 = 0.9936$$
$$Combined\ reliability_C = 1 - 0.01 = 0.99$$

$$System\ Reliability = Combined\ reliability_A Combined\ reliability_B Combined\ reliability_C$$
$$System\ Reliability = (0.9775)(0.9936)(0.99) = 0.9615 = 96.15\%$$

It is clear from the results, that running all the machines increases the reliability by 25.77%, which is a large increase. It would thus be recommended to management that Magnaplex should operate both of machines A, B and C in parallel to ensure a much higher reliability.

Delivery Vehicles

For the delivery process, there are 20 delivery vehicles available, of which 19 is required to be operating at any time to give reliable service. During the past 1560 days, the number of days that there was only 20 vehicles available was 190 days, only 19 vehicles available was 22 days, only 18 vehicles available was 3 days and 17 vehicles available only once.

It is given that there are 20 delivery vehicles available, of which 19 are required to be operating at any given time.

For the vehicles, the probability of having all vehicles available is 0.86169, meaning that for a year, there will be roughly 314 days that all vehicles are available.

The probability of having 19 vehicles available is 0.128726, meaning that for a year, there will be roughly 47 days that 20 vehicles are available.

In order for the service to be reliable, at least 19 out of the 20 vehicles need to operating at all times. Therefore, the total reliability is 0.86168 plus 0.128762 which gives roughly 0.99, meaning that 361 days in a year will have at least 19 vehicles operating.

Drivers are needed to operate the vehicles. There are 21 drivers in total, and the probability of having all 21 drivers is 0.9343, meaning that for 341 days of a year, there are 21 drivers available. The probability of having 20 drivers available is 0.0659, meaning that for a year, there are 23 days where 20 drivers are available.

Thus the total reliability for the drivers is 0.996, meaning that for 364 days in a year at least 20 drivers are available.

By combining the reliability of the vehicles and the driver, this gives a combined total reliability of 0.9883, meaning that for 361 days in a year, service delivery is reliable.

If the number of vehicles is increased by one to 21, the total reliability becomes 0.9995, which translates to roughly 364 days where service delivery is reliable.

References

Radečić, D. (2022). *MANOVA in R – How To Implement and Interpret One-Way MANOVA | R-bloggers*. [online] Available at: https://www.r-bloggers.com/2022/01/manova-in-r-how-to-implement-and-interpret-one-way-manova/ [Accessed 16 Oct. 2022].

Hessing, Ted. (2014). Statistical Process Control (SPC). *Six Sigma Study Guide* [online] Available at: https://sixsigmastudyguide.com/statistical-process-control-spc/. [Accessed 11 Oct. 2022].

Berardinelli, C. (2018). *A Guide to Control Charts | iSixSigma*. [online] Isixsigma.com. Available at: https://www.isixsigma.com/tools-templates/control-charts/a-guide-to-control-charts/ [Accessed 9 Oct. 2022].