QUALITY ASSURANCE 344
ECSA PROJECT

Lawrence Wiid - 21682909

21/10//2022

Contents

List of figures

List of tables

Introduction

An analysis of sales data for an online business needs to be completed. The client sales data indicates various variables with regards to the sales and these variables will help to perform an analysis on the data which will provide company management with valuable information and recommendations to improve their process if possible.

The analysis will include data wrangling in order to remove compromised or incomplete values, furthermore descriptive statistics will be used to understand the data and statistical control measures will be implemented to measure the capability of the processes. The process will also include optimizing delivery times and calculating the reliability of the delivery process.

Part 1: Data Wrangling

The data wrangling involved separating the sales data into a valid data set and an invalid data set. The data was classified under the invalid data set if there were any NA values or negative values, thus these entries were removed from the original sales data to create the invalid data set as well as creating new valid data set with the remaining values.

The invalid data set had numerous anomalies which include 22 invalid entries under the pricing feature, 17 of which were missing values and further five values were negative.

Invalid Data

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

Figure 1.1: Invalid Data Set

Ordered Valid Data

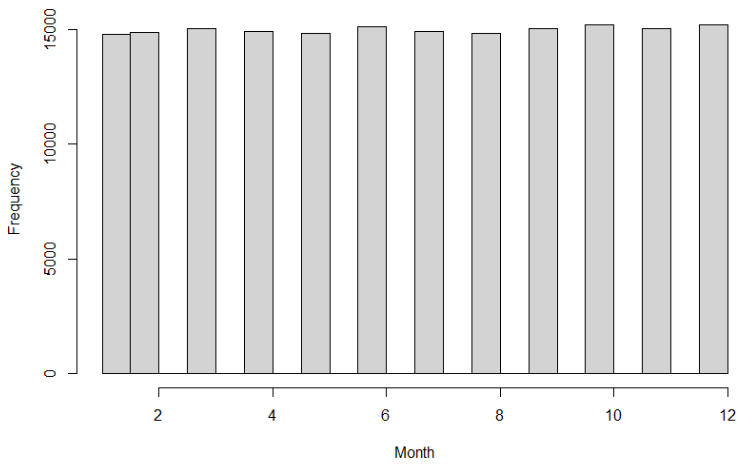
	Primary.Key	Secondary Key	ID	AGE	Class	Price	Year	Month	Day	Delivery.time	Why.Bought
463	1	463	47101	50	Clothing	1030.86	2021	1	1	9.0	Recommended
2627	2	2627	88087	21	Clothing	428.03	2021	1	1	10.0	Recommended
3374	3	3374	25418	68	Household	13184.41	2021	1	1	48.5	Website
5288	4	5288	13566	94	Household	7021.90	2021	1	1	42.0	Recommended
8182	5	8182	84692	35	Clothing	475.18	2021	1	1	9.0	Recommended
9272	6	9272	46305	72	Clothing	580.98	2021	1	1	8.5	Random
9712	7	9712	92105	45	Household	6877.00	2021	1	1	43.0	Recommended
12163	8	12163	21614	27	Clothing	513.13	2021	1	1	9.5	Recommended
12195	9	12195	12174	56	Household	14538.64	2021	1	1	41.5	EMail
20004	10	20004	84558	74	Food	255.41	2021	1	1	2.0	Recommended
20509	11	20509	15630	32	Clothing	164.56	2021	1	1	9.0	Recommended
21970	12	21970	81216	87	Clothing	173.76	2021	1	1	10.0	Recommended
27161	13	27161	56240	45	Household	17681.94	2021	1	1	45.5	Website
27638	14	27638	24396	30	Clothing	1018.21	2021	1	1	8.5	Recommended
30778	15	30778	12235	28	Technology	21096.86	2021	1	1	15.0	Website
34277	16	34277	30290	43	Household	10573.67	2021	1	1	51.0	Recommended
34950	17	34950	40035	77	Household	16548.61	2021	1	1	51.5	Recommended
35153	18	35153	36435	53	Technology	23304.75	2021	1	1	14.0	Browsing
37187	19	37187	49974	67	Sweets	332.46	2021	1	1	2.5	Recommended
42139	20	42139	36292	75	Food	205.96	2021	1	1	3.0	Recommended

Figure 1.2: Ordered Valid Data Set

The valid data set has been ordered according to date, from the earliest date to the most recent. A primary key has also been inserted for additional information.

Part 2: Descriptive statistics

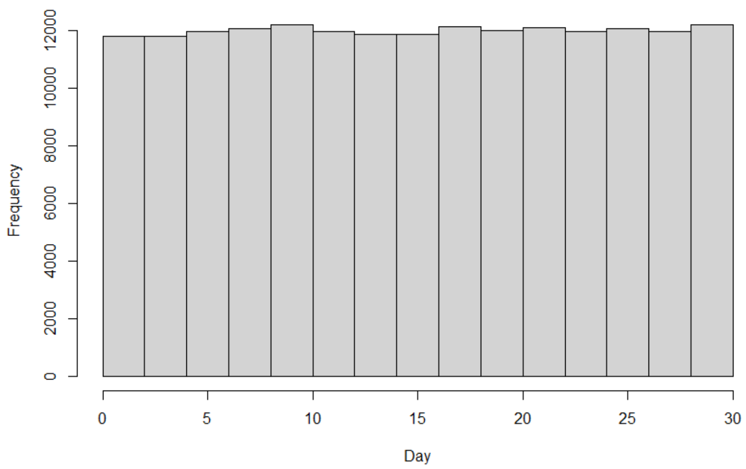
Frequency of all items sold in each month



The distribution has little to no variation as the amount of items sold are very similar every month. Thus, the month that the items are sold in doesn't affect the amount of items sold.

Figure 2.1: Histogram showing the frequency of items sold in each month

Frequency of all items sold on each day of a month



The distribution also has little no to variation as the amount of items sold are very similar for each day of the month. Thus, the day of the month in which an item is sold does not affect the amount of items sold.

Figure 2.2: Histogram showing the frequency of items sold on each day of a month

Amount of items sold each year for the various classes

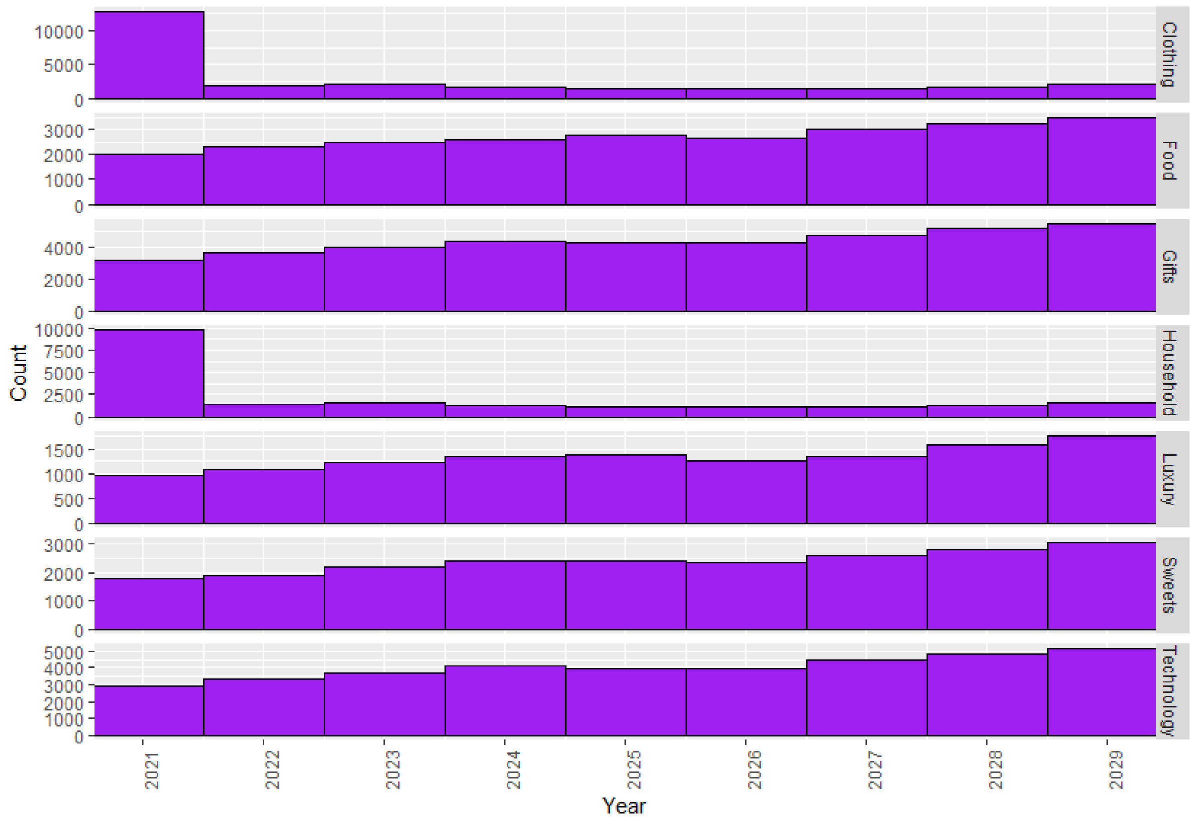
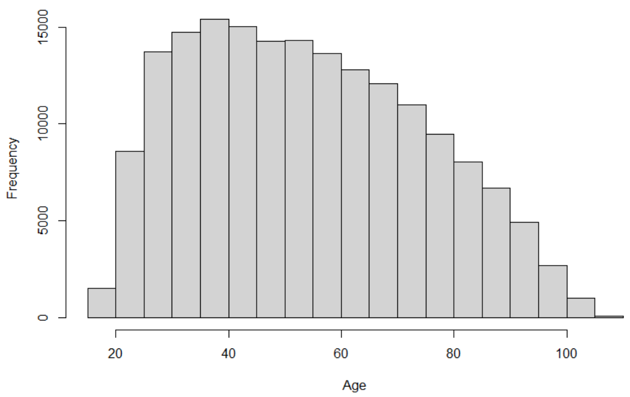


Figure 2.3: Histogram showing the count per class for each year

The distribution from the histogram above indicates a slight increase in sales each year for food, gifts, luxury, sweets and technology items, however in 2026 there is a slight decrease in sales for all these classes. For clothing and household items, both classes had large sales amounts in 2021 compared to the following years as the sales dropped considerably for the years 2022 to 2029.

Age for all instances



The distribution of age for all instances follows a distribution which is right skewed, indicating that the majority of the people are between the ages of 25 and 70 and the frequency of age becomes less as the age increases past 55. The distribution also indicates that there a few people under 25 making purchases.

Figure 2.4: Histogram showing the age

Age vs Class

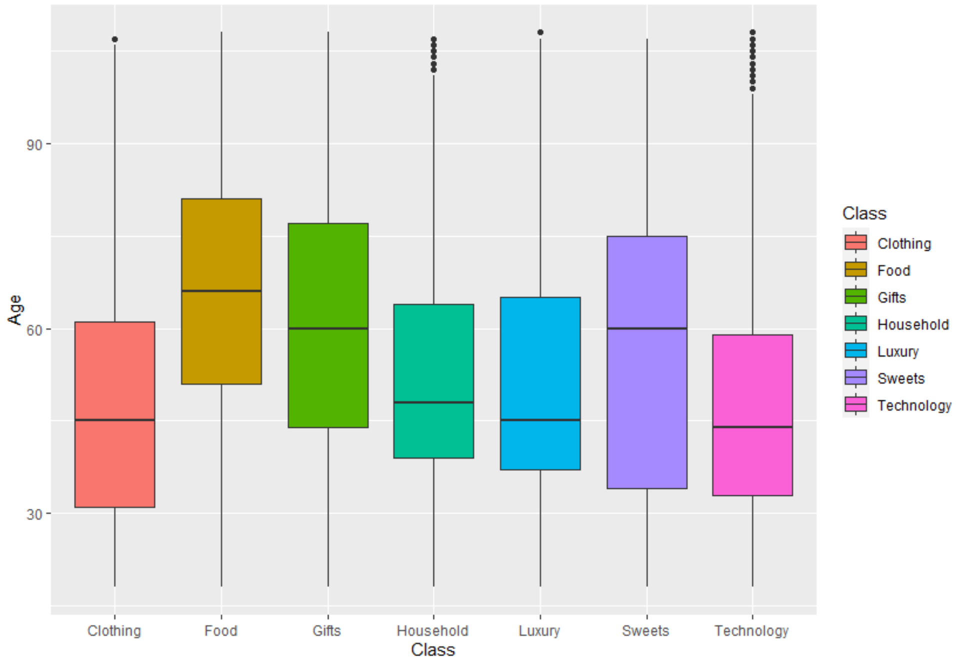
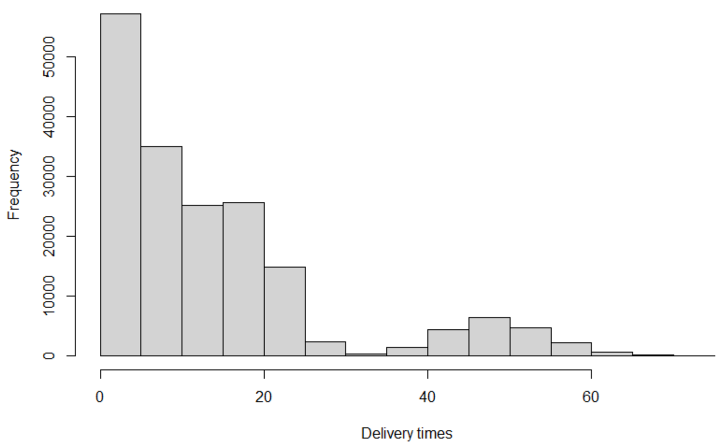


Figure 2.5: Box plot showing the age vs class

The boxplot above indicates that people who buy clothing, technology and luxury items tend to be of a younger age, which is logical as the younger generation tend to have more of an interest in such items. However there is a fairly large age range away from the median for all the classes. The boxplots also indicates that people who buy food, gifts and sweets tend to be of older age, above 60 and household items being somewhere in the middle of the other classes.

Delivery times for all instances



The distribution of delivery times for all instances is right skewed with a small increase in delivery times between 40 and 60. The majority of the delivery occur within 0 to 20 hours, however a small amount of very long delivery times are driving up the mean.

Figure 2.6: Histogram showing the delivery times

Delivery time vs Class

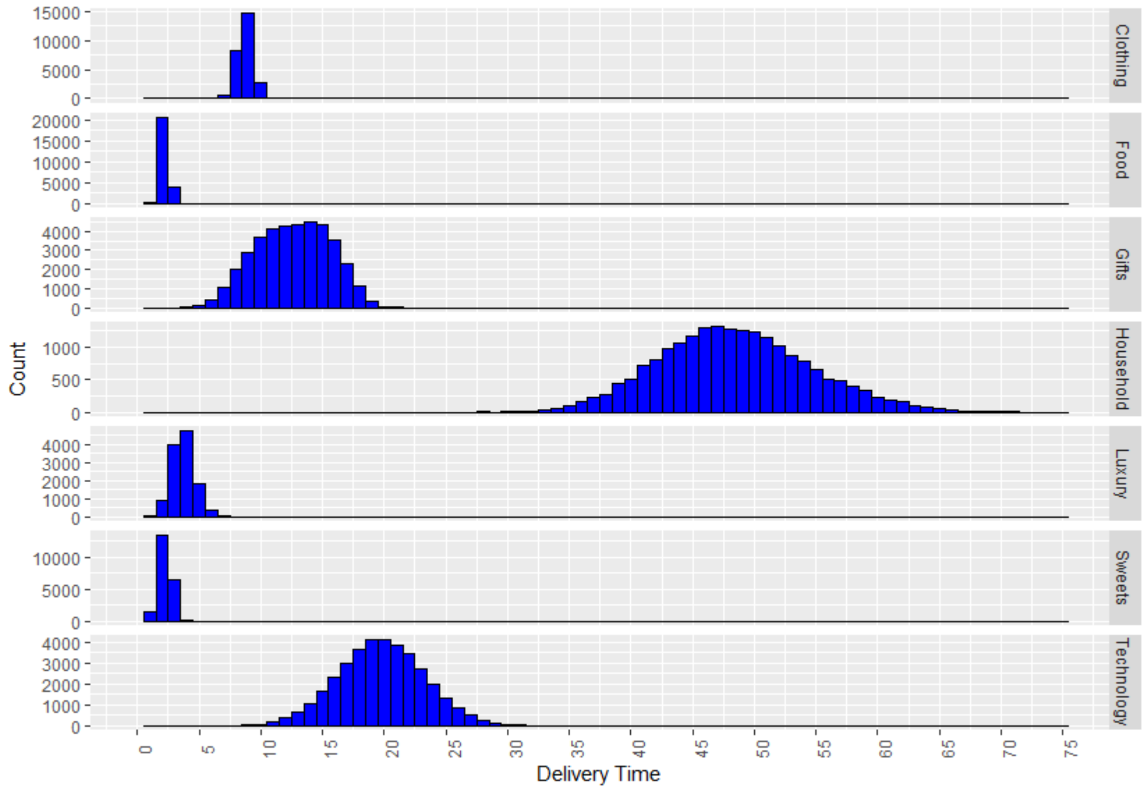


Figure 2.7: Histogram showing the delivery time vs class

The distributions for delivery time for the technology, household, gifts and luxury items all follow a normal distribution, where household items take the longest and food and sweets have the shortest duration, which is logical when considering the type of items purchased. Clothing, food and sweets all have a small range in delivery times while household items has the largest range of delivery time length.

Price vs Class

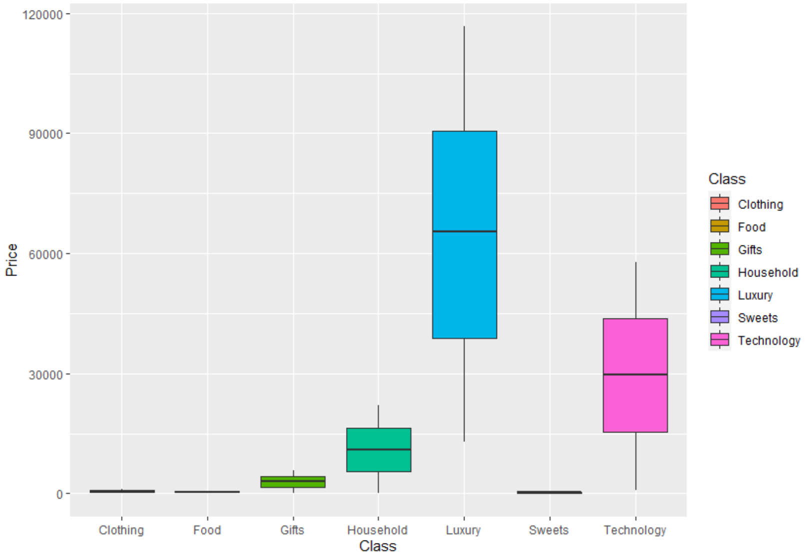


Figure 2.8: Box plot showing the price vs class

Shown above in figure 2.8 are the classes which are least expensive and most expensive. As shown, the sweets, food and clothing classes have the lowest average prices and luxury items have the most expensive item and most expensive average price. This is to be expected as food, sweets and clothing are generally less expensive than the other class items such as luxury or technology items and accordingly, the technology items has the second highest price average. The third most expensive average price is for household items followed by gifts. These results are to be expected and are not out of the ordinary.

Why Bought?

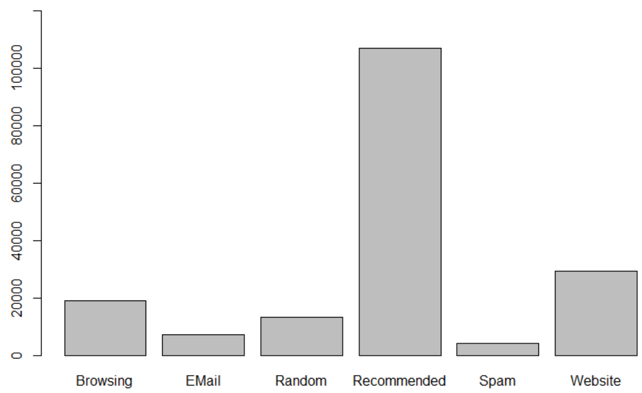


Figure 2.9: Histogram showing the why bought

From the histogram in figure 2.9, the most popular reason why people bought the products was due to the products being recommended. This is logical as people would be recommended products based on previous purchases and would thus be most likely to purchase the recommended product. The second and third most popular reason was through browsing and websites as many people probably have access to the internet. Spam and email are the least popular reason why people purchased the products which is understandable as most people ignore such emails or spam.

Process Capability indices

The process capability is the ability of a process to meet specifications, in this case, the delivery times for technology class items will be evaluated. The upper specification limit (USL) = 24 hours with a lower specification limit (LSL) = 0. An LSL of 0 is logical because these specifications are intended for the delivery times and you cannot have a delivery time less than 0

Table 2.1: Process capability indices

Mean =	20.01095
Std dev =	3.501993
Cp =	1.142
Cpu =	0.379
Cpl =	1.904
Cpk =	0.379

The Cp indicates whether a process is capable of meeting specifications and the Cpk also indicates whether the process is cable of meeting specifications along with the ability to adhere to the target specification. A Cpk of 1 is considered to be marginally capable and a Cpk of 1.3 is considered to be capable, values higher than 1.3 translates to higher capability.

By looking at table 2.1, a Cp value of 1.142 suggests that the process has the potential to be capable, however the Cpk value is 0.379, less than one, suggesting that the process is currently not capable and not centred. Thus, fundamental changes will have to be made in order to improve the process and reduce variation and spread within the delivery times for technology items.

Part 3: Statistical Process Control

3.1 initial SPC control charts

By looking at the first 30 samples of size 15 per sample of the delivery process times, Statistical Process Control (SPC) analyses whether the process is in control. Thus, by initialising X&s charts for every class of sales we can then determine by graphical means if these samples are within the acceptable range. The s-charts will test whether the process variation is in control and if there are any s-charts with out of control patterns, then we will have to remove the samples that are beyond the control limits first in order to produce accurate x-Bar charts to be analysed.

In the figures below the red lines represent the upper and lower control limits while the blue line represents the centre line.

Clothing

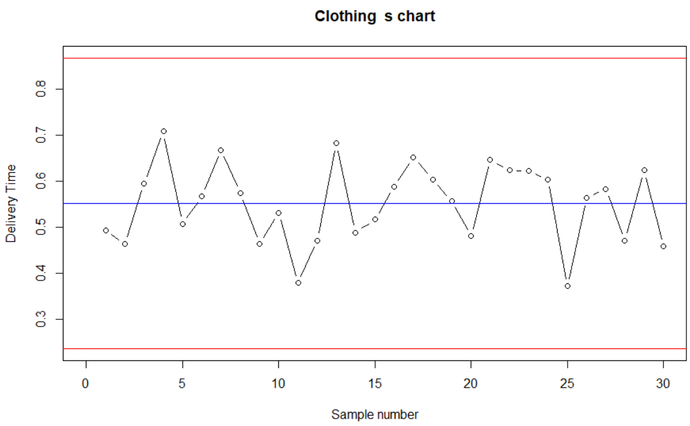


Figure 3.1.1: initial clothing s-chart

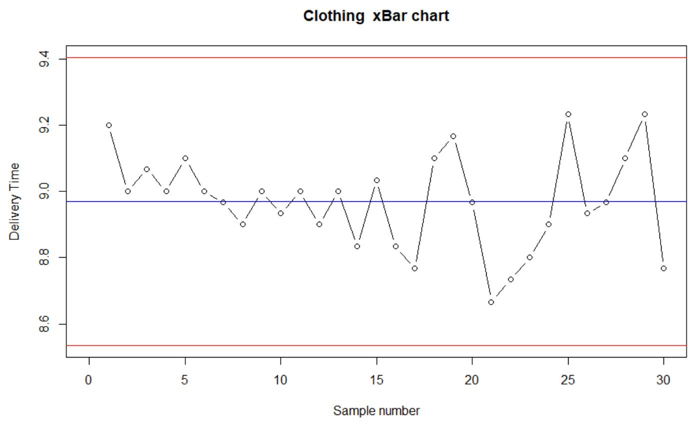


Figure 3.1.2: initial clothing x-Bar chart

From the s-chart above (figure 3.1.1) there are no sample points that are beyond the control limits, meaning that the s-chart is in control and the x-Bar chart will have accurate control limits. The sample points in the clothing x-Bar chart (figure 3.1.2) all fall within the control limits, indicating that the initial clothing samples are in statistical control.

Household

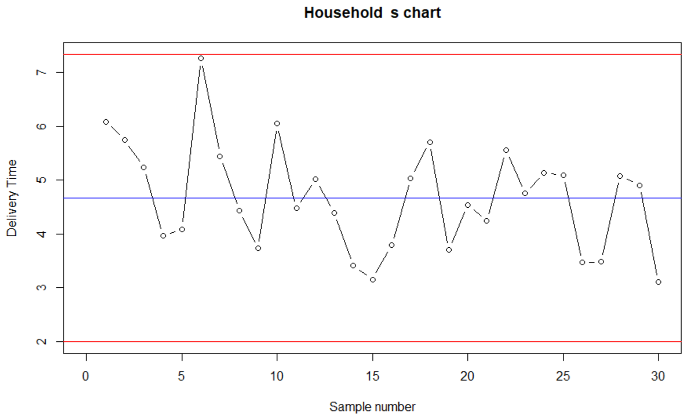


Figure 3.1.3: initial household s-chart

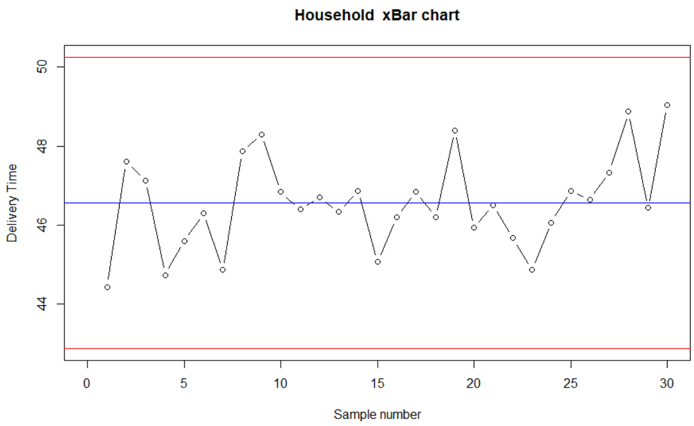


Figure 3.1.4: initial household x-Bar chart

From the s-chart above (figure 3.1.3) there are no sample points that are beyond the control limits, meaning that the s-chart is in control and the x-Bar chart will have accurate control limits. The sample points in the household x-Bar chart (figure 3.1.4) all fall within the control limits, indicating that the initial household samples are in statistical control.

Food

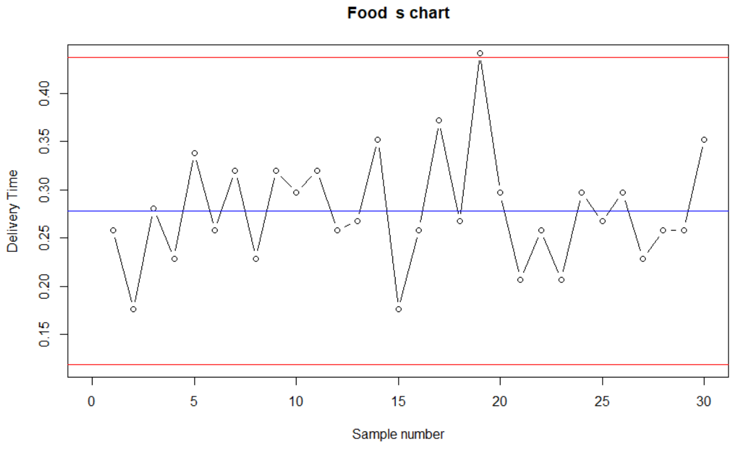


Figure 3.1.5: initial food s-chart

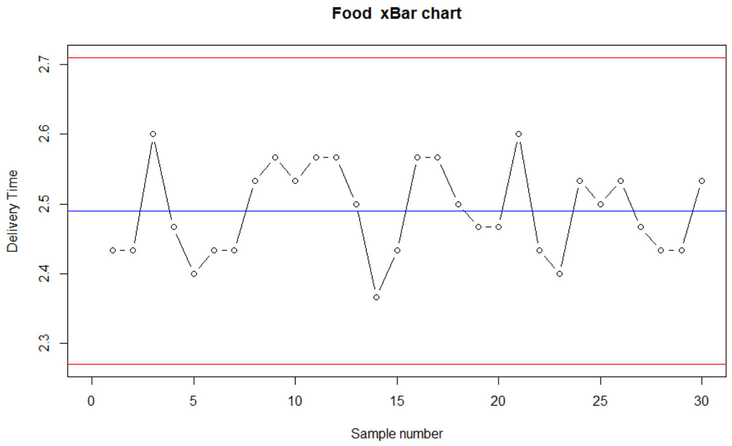


Figure 3.1.6: initial food x-Bar chart

From the food s-chart above (figure 3.1.5) sample point 19 is beyond the control limit, meaning that the sample needs to be removed in order to calculate the new control limits to produce an accurate x-Bar chart.

Technology

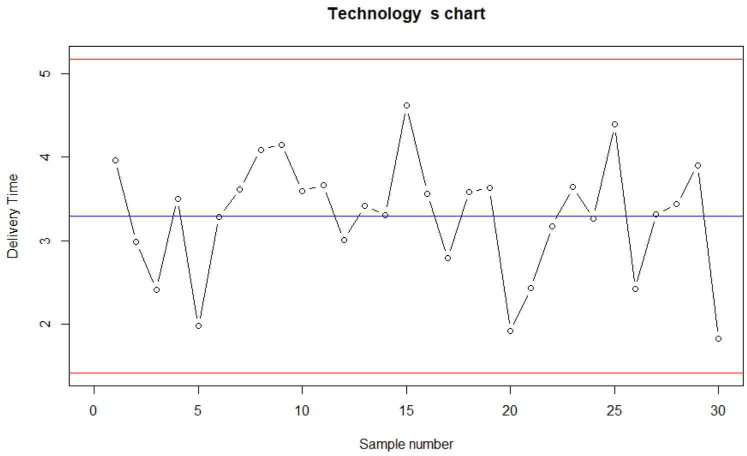


Figure 3.1.7: initial technology s-chart

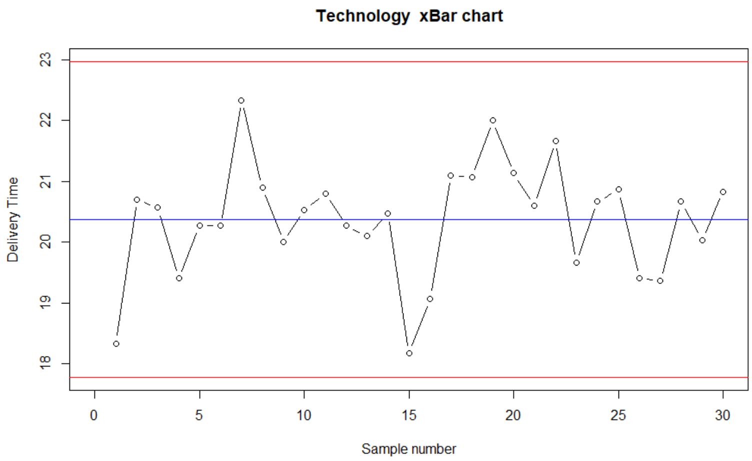


Figure 3.1.8: initial technology x-Bar chart

From the s-chart above (figure 3.1.7) there are no sample points that are beyond the control limits, meaning that the s-chart is in control and the x-Bar chart will have accurate control limits. The sample points in the technology x-Bar chart (figure 3.1.8) all fall within the control limits, indicating that the initial technology samples are in statistical control.

Sweets

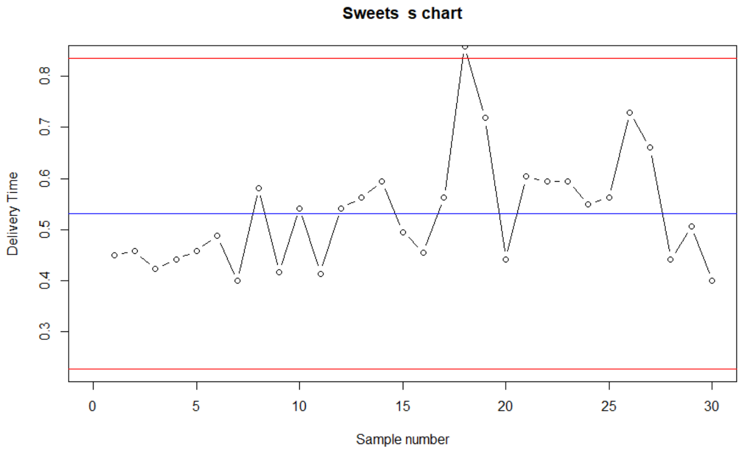


Figure 3.1.9: initial sweets s chart

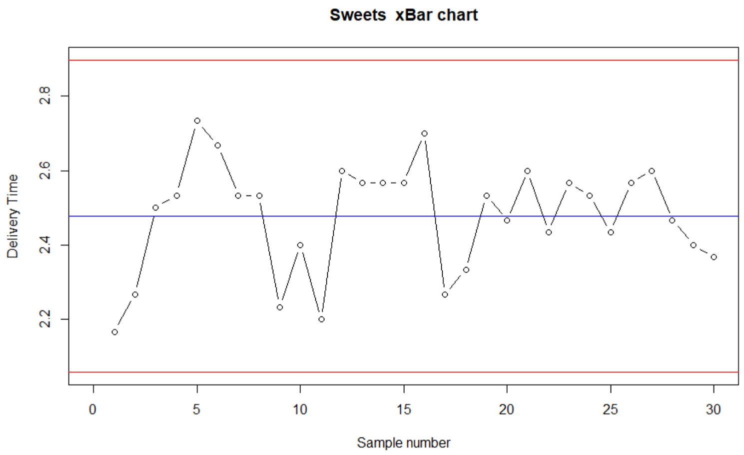


Figure 3.1.10: initial sweets x-Bar chart

From the sweets s-chart above (figure 3.1.9) sample point 18 is beyond the control limit, meaning that the sample needs to be removed in order to calculate the new control limits to produce an accurate x-Bar chart.

Gifts

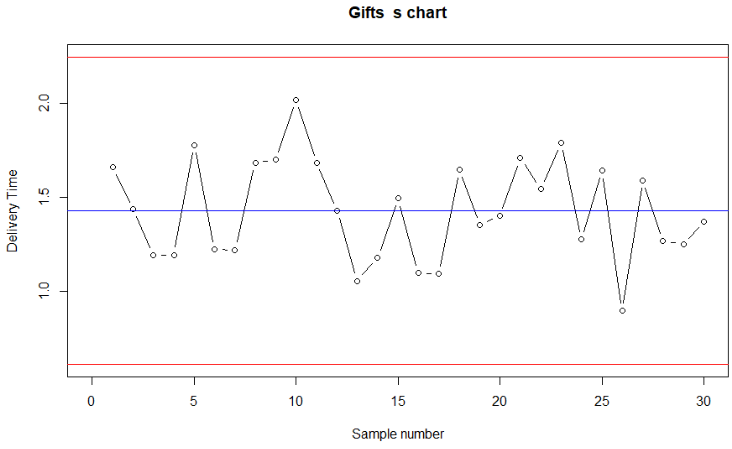


Figure 3.1.11: initial gifts s-chart

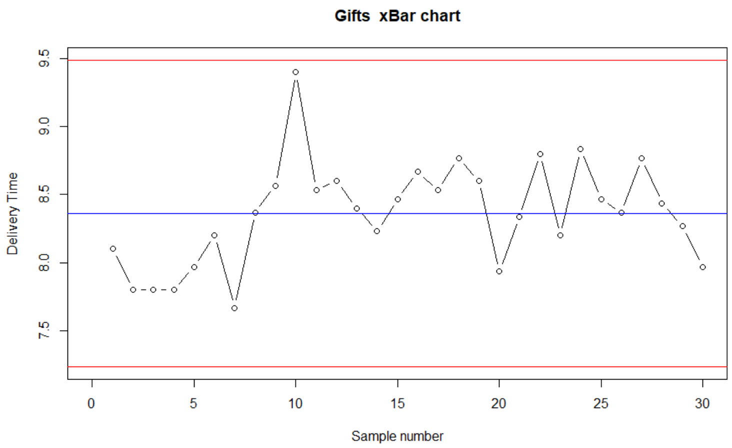


Figure 3.1.12: initial gifts x-Bar chart

From the s-chart above (figure 3.1.11) there are no sample points that are beyond the control limits, meaning that the s-chart is in control and the x-Bar chart will have accurate control limits. The sample points in the gifts x-Bar chart (figure 3.1.12) all fall within the control limits, indicating that the initial gifts samples are in statistical control.

Luxury

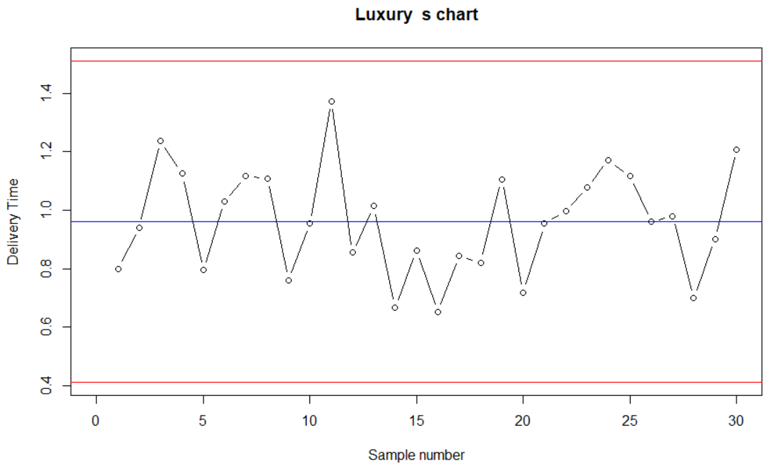


Figure 3.1.13: initial luxury s-chart

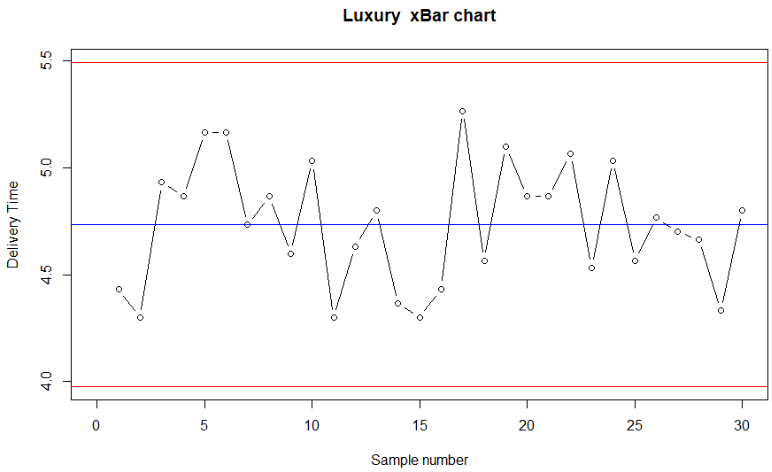


Figure 3.1.14: initial luxury chart x-Bar chart

From the s-chart above (figure 3.1.13) there are no sample points that are beyond the control limits, meaning that the s-chart is in control and the x-Bar chart will have accurate control limits. The sample points in the luxury x-Bar chart (figure 3.1.14) all fall within the control limits, indicating that the initial luxury samples are in statistical control.

New initial s-charts for food and sweets with outliers removed

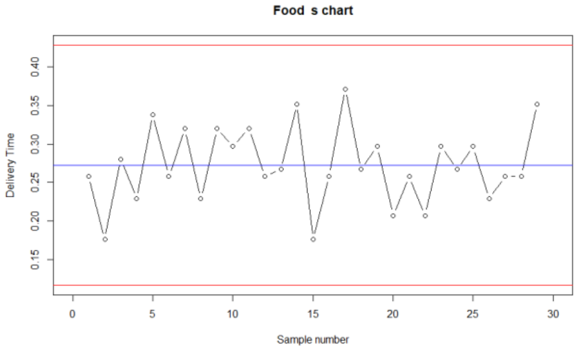


Figure 3.1.15: new food s-chart

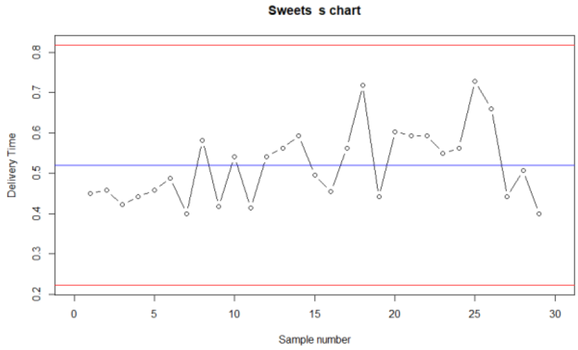


Figure 3.1.16: new sweets s-chart

For the initial samples of food and sweets the outliers have been removed and the delivery time process in now in statistical control.

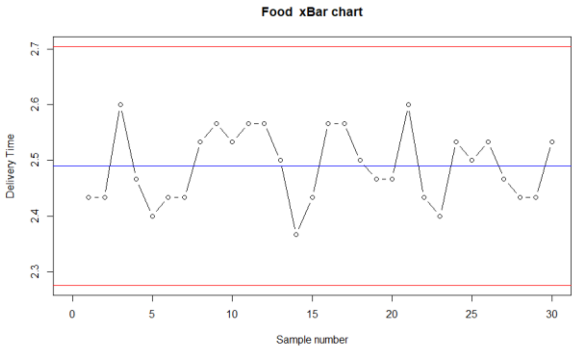


Figure 3.1.17: new food x-Bar chart

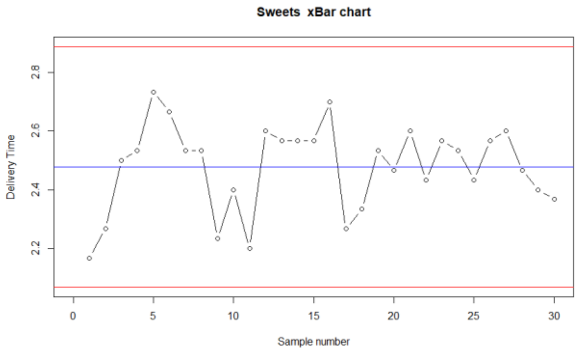


Figure 3.1.18: new sweets x-Bar chart

Tabled values for s chart

Table 3.1: s-chart table values with new control limits

Class	UCL	U2Sigma	U1Sigma	CL	L1Sigma	L2Sigma	LCL
Clothing	0.866559568463719	0.761455227250562	0.656350886037405	0.551246544824249	0.446142203611092	0.341037862397935	0.235933521184778
Household	7.34418006586244	6.45341013420991	5.56264020255739	4.67187027090486	3.78110033925233	2.89033040759981	1.99956047594728
Food	0.428372325526015	0.376415376798093	0.324458428070171	0.272501479342249	0.220544530614327	0.168587581886404	0.116630633158482
Technology	5.18056970372824	4.55222240293678	3.92387510214531	3.29552780135385	2.66718050056238	2.03883319977091	1.41048589897945
Sweets	0.817573461203726	0.71841060713067	0.619247753057614	0.520084898984558	0.420922044911502	0.321759190838447	0.222596336765391
Gifts	2.24633333311156	1.9738772969496	1.70142126078763	1.42896522462567	1.15650918846371	0.884053152301749	0.611597116139788
Luxury	1.51105176847233	1.32777746576534	1.14450316305835	0.961228860351357	0.777954557644365	0.594680254937373	0.411405952230381

Tabled values for x-Bar chart

Table 3.2: x-Bar chart table values

Class	UCL	U2Sigma	U1Sigma	CL	L1Sigma	L2Sigma	LCL
Clothing	9.40493352386633	9.25995568257756	9.11497784128878	8.97	8.82502215871122	8.68004431742245	8.53506647613367
Household	50.2483278659662	49.0196259847182	47.7909241034702	46.5622222222222	45.3335203409742	44.1048184597263	42.8761165784783
Food	2.70500366720103	2.63333577813402	2.56166788906701	2.49	2.41833211093299	2.34666422186598	2.27499633279897
Technology	22.9746158797126	22.1078920679566	21.2411682562005	20.3744444444444	19.5077206326884	18.6409968209323	17.7742730091763
Sweets	2.88812476307659	2.75134243464366	2.61456010621072	2.47777777777778	2.34099544934484	2.2042131209119	2.06743079247896
Gifts	9.48856467334077	9.11274681926422	8.73692896518766	8.36111111111111	7.98529325703456	7.60947540295801	7.23365754888145
Luxury	5.49396512637278	5.24116193610037	4.98835874582796	4.73555555555556	4.48275236528315	4.22994917501074	3.97714598473833

3.2 Continued SPC for all samples

In section 3.1 the X&s charts only included the first 30 samples, and were altered to be in statistical control. In this section, the x-Bar charts include all the samples for each class, thus being able to see whether the delivery process remains in control or goes out of control.

Clothing

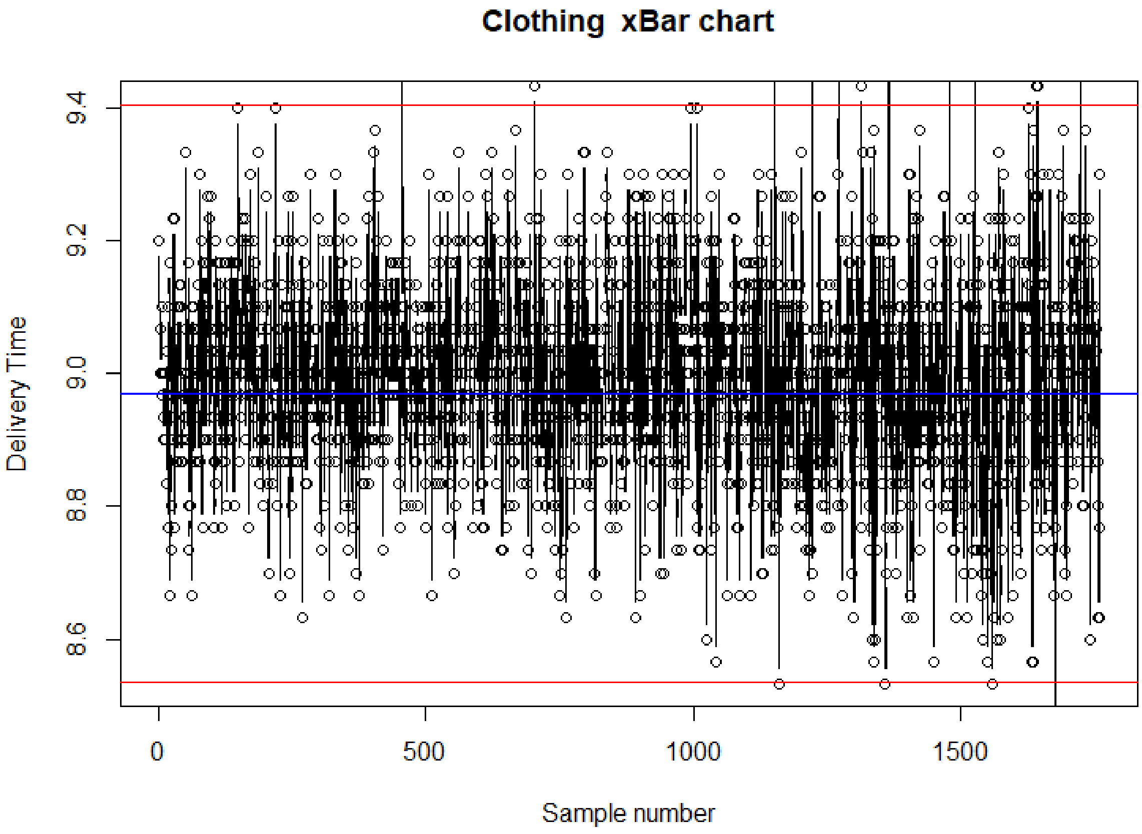


Figure 3.2.1: clothing x-Bar chart

The clothing delivery time process has stayed in control, there a very few instances where the sample points fall outside of the control limits. There is also a similar amount of sample points above and below the centre line, indicating that the sample means follows a normal distribution.

Household

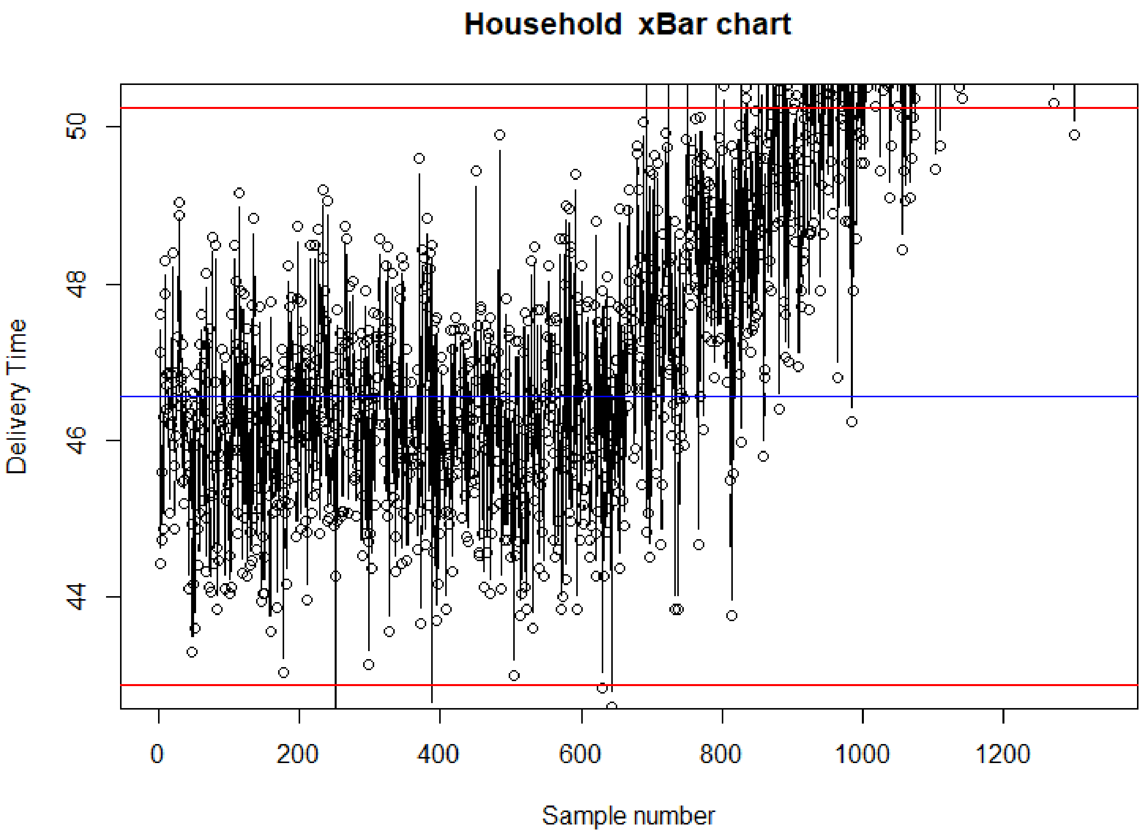


Figure 3.2.2: household x-Bar chart

The household delivery time has gone out of control, initially up until sample number 600 the chart stayed in control but then afterwards had an increasing trend beyond the control limits. This indicates that the delivery time for household items became longer over time, perhaps due to people purchasing the items from further locations as the business became more popular.

Food

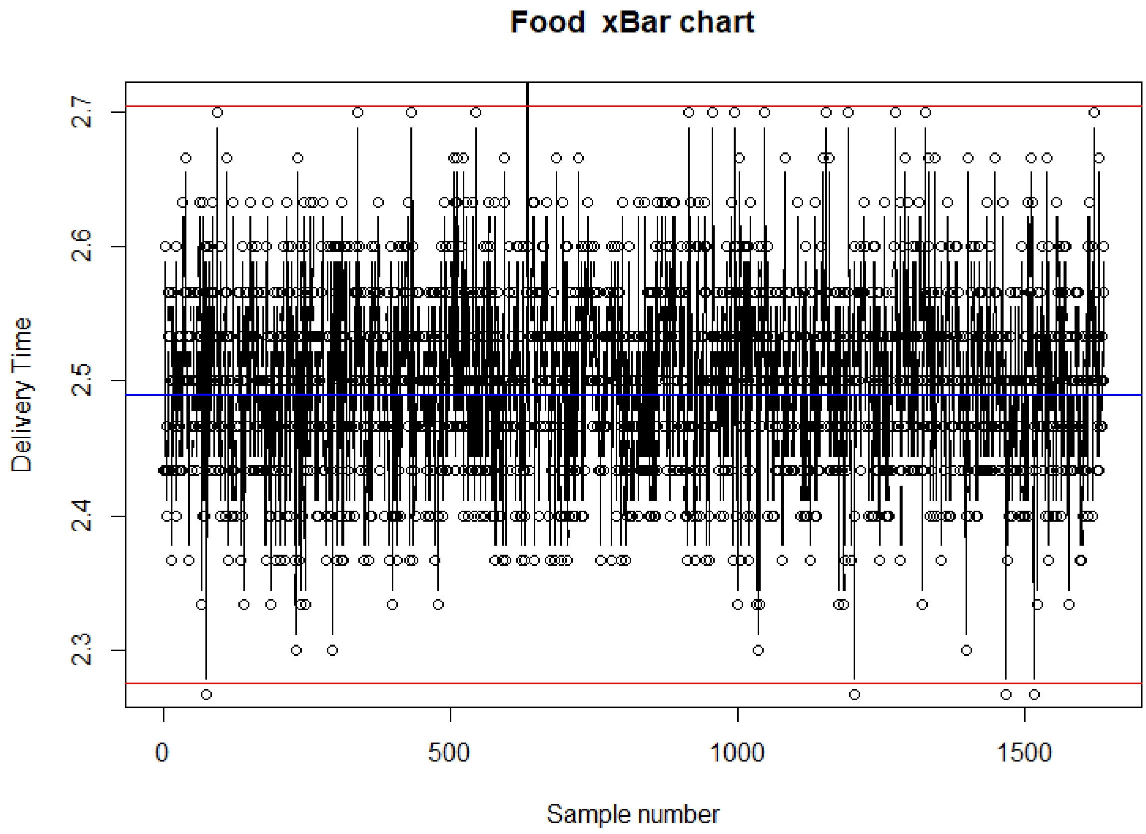


Figure 3.2.3: food x-Bar chart

The food delivery time process has stayed in control, there a very few instances where the sample points fall outside of the control limits. There is also a similar amount of sample points above and below the centre line, indicating that the sample means follows a normal distribution.

Technology

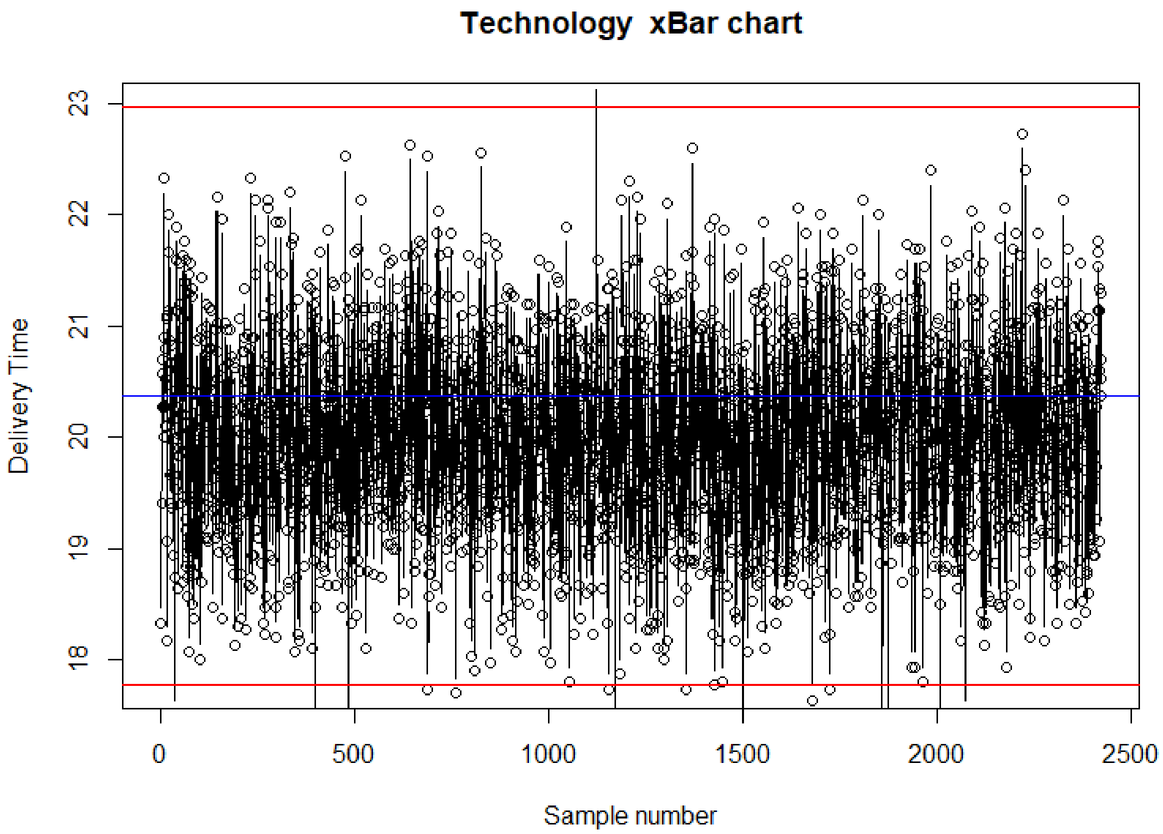


Figure 3.2.4: technology x-Bar chart

The technology delivery time process has stayed in control, there a very few instances where the sample points fall outside of the control limits. There is also a similar amount of sample points above and below the centre line, indicating that the sample means follows a normal distribution.

Sweets

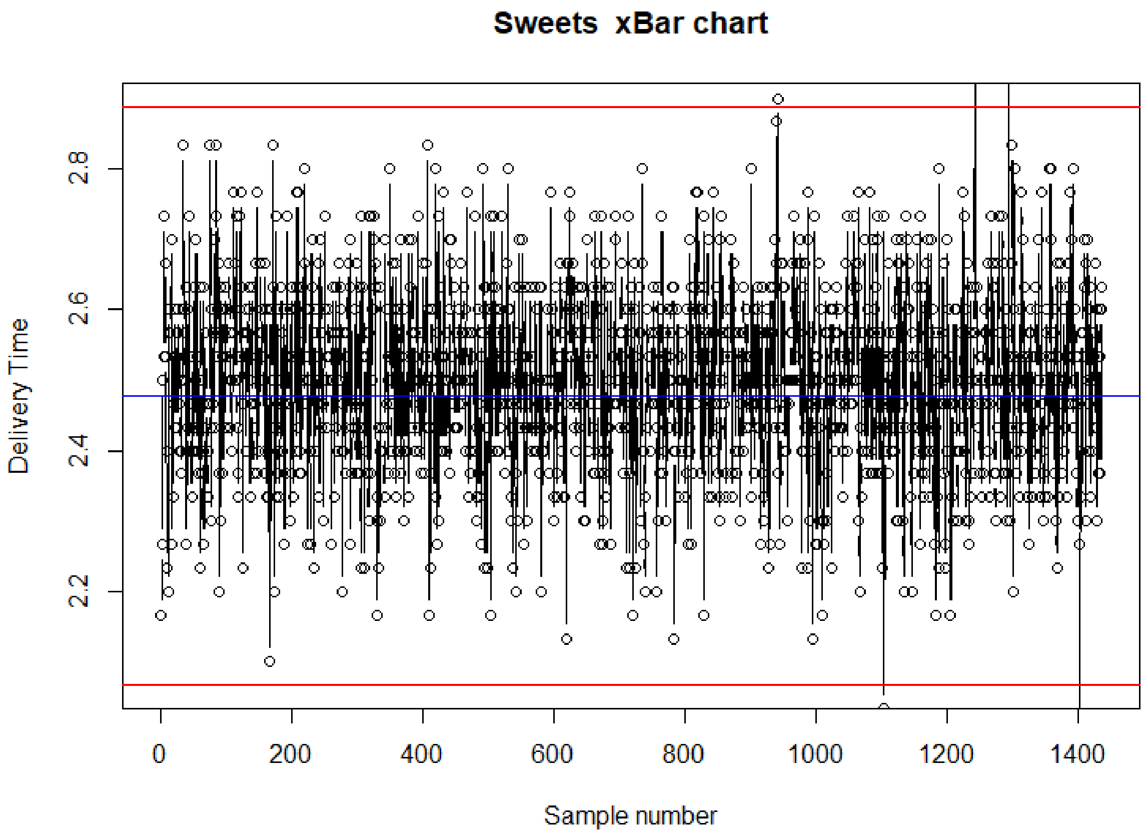


Figure 3.2.5: sweets x-Bar chart

The sweets delivery time process has stayed in control, there a very few instances where the sample points fall outside of the control limits. There is also a similar amount of sample points above and below the centre line, indicating that the sample means follows a normal distribution.

Gifts

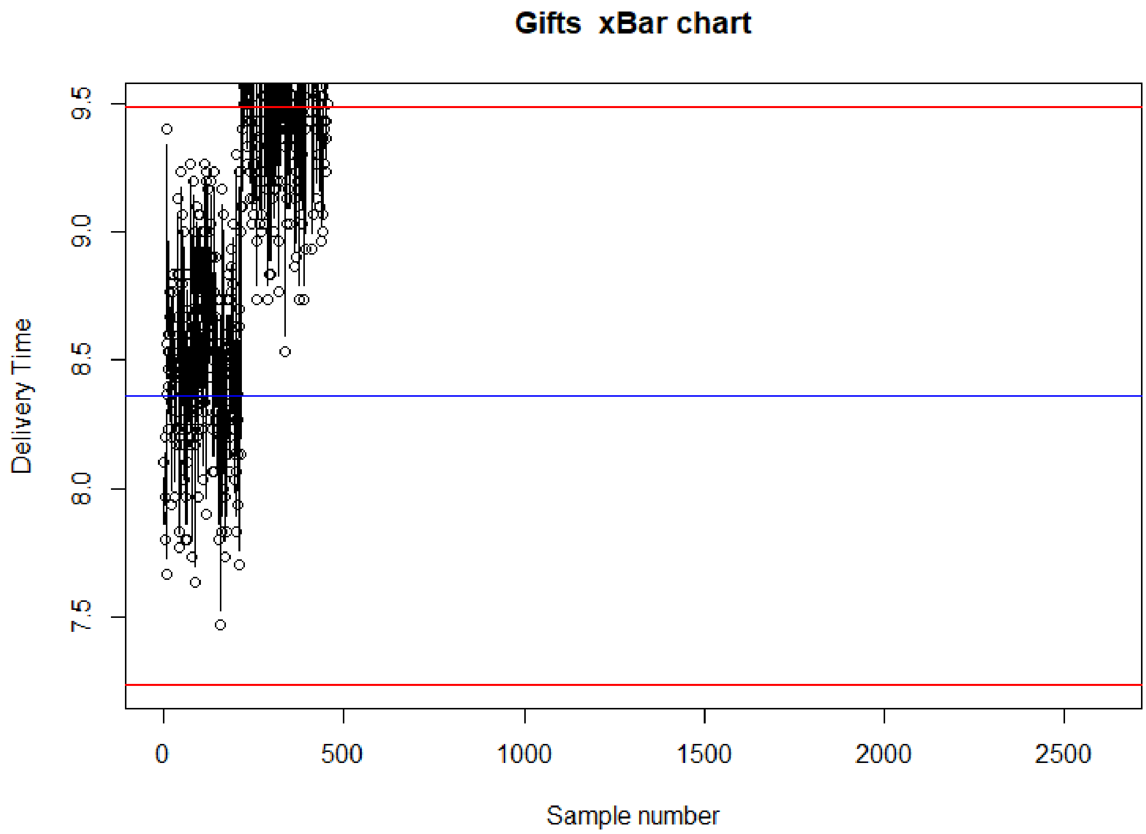


Figure 3.2.6: gifts x-Bar chart

The gifts delivery time has gone out of control, initially for the first 250 sample points the chart stayed in control but then afterwards had a drastic increase beyond the control limits. This indicates that the delivery time for gift items became much longer very quickly, perhaps due to an increase in sales while having limited amounts of stock available.

Luxury

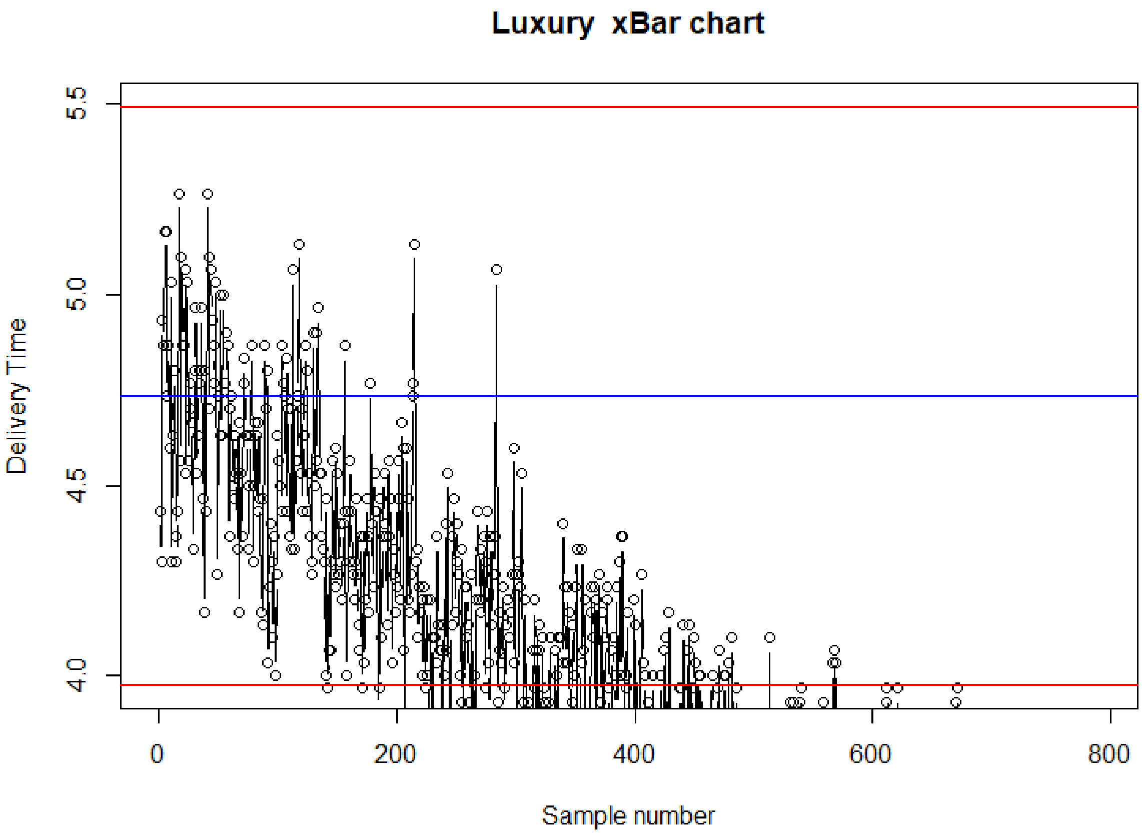


Figure 3.2.7: luxury x-Bar chart

The luxury delivery time has gone out of control, initially for the first 200 sample points the chart stayed in control even though the sample points had a decreasing trend moving below the centre line. However after the first 200 samples the delivery times continued to decrease beyond the control limits. This indicates that the delivery time for gift items continued to become shorter over a period of time, perhaps due to a new improving delivery process for luxury items.

Part 4: Optimising the delivery processes

4.1 outside control limits

Rule A: x-Bar outside of the outer control limits

Table 4.1: Rule A

Class	Number of outliers	First	second	Third	Third last	Second last	Last
Clothing	17	455	702	1152	1677	1723	1724
Household	400	252	387	629	1335	1336	1337
Food	5	75	633	1203	1467	1515	NA
Technology	17	37	398	483	1872	2009	2071
Sweets	5	942	1104	1243	1294	NA	1403
Gifts	2290	213	216	218	2607	2608	2609
Luxury	434	142	171	184	789	790	791

Rule B: The most consecutive samples of sample standard deviations between -0.3 and +0.4 sigma-control limits

4.2 Type 1 error probability

A type I error occurs when the process is wrongly said to be out of control when it is actually in control, also known as a manufacturer's error.

The type I error probability for rule A is = 0.002699796
Rule A thus has a low probability of being incorrectly judged as out of control

The type I error probability for rule B is = 0.8224429
Rule B has a high chance of being incorrectly judged as out of control, most likely due to the -0.3 and +0.4 adjusted control limits

4.3 Best delivery time to maximize profit

Based on the delivery time for technology items, we need to determine by how many hours to reduce the average delivery time to maximise the profits. The company loses R329 per item per late hour and costs R2.5 per item per hour to reduce the average by one hour. A delivery is considered to be late if the delivery time is more than 26 hours. Thus we need to reduce the cost as much as possible in order to maximize profits.

First we use the current delivery times to determine the current penalty costs for comparison purposes.

The total cost for late deliveries are currently = R758 674

By reducing the delivery time by x hours for every item bought up until 10 hours the total costs are as follows

Table 4.2: total cost for x hours reduced

X hours reduced	Total cost
1	460992.5
2	250002.5
3	273918.5
4	363470.0
5	454337.5
6	545205.0
7	636072.5
8	726940.0
9	817807.5
10	908675.0

Thus by looking above in table 4.2 it is clear that reducing the delivery time average by 2 hours will cost the least and will thus, maximize profits.

What is seen in the process above in table 4.2 is similar to that of a Taguchi loss function as it follows a quadratic equation shape, resulting in higher costs as you move away from the desired target.

4.4 Type II error probability

Type II errors occur when the process is unstable or out of control but the graphs give no indication of this, thus type II errors occur when we presume the process is in control when it is actually out of control, also known as a consumer's error.

The likelihood of making a type II error for 4.1 rule A:

For technology items the UCL = 22.9746 and LCL = 17.7743 stays the same as in table 3.2, however the delivery process average has been moved to 23 hours. Thus the probability of a type II error, which is the probability that the samples remain within the control limits after the delivery process average has been moved is = 0.4928718

Part 5: DOE and MANOVA

For the MONOVA test an alpha value of 0.01 was chosen and thus a test will be significant if $p < 0.01$.

```
 Response months :
                Df  Sum Sq Mean Sq F value Pr(>F)
class            6      87  14.576  1.2219 0.2913
Residuals   179971 2146871  11.929

 Response days :
                Df   Sum Sq Mean Sq F value Pr(>F)
class            6      668 111.302   1.488 0.1777
Residuals   179971 13461680  74.799
```

Figure 5.1: MANOVA for day and month vs class

The figure above (figure 5.1) shows that the day and month a sale was made is not significantly different among the class of item sold

```
 Response ages :
                Df   Sum Sq Mean Sq F value    Pr(>F)
class            6  8422401 1403733    3805 < 2.2e-16 ***
Residuals   179971 66394669     369
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

 Response prices :
                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

 Response years :
                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 deliverys :
                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
```

Figure 5.2: MANOVA for age, price, year and delivery time vs class

The figure above (figure 5.2) shows that the age, price, year and delivery time of a sale that was made was significantly different among the class of item sold with delivery time clearly having the largest F value thus having the strongest relationship. The weakest relationship was with years and then followed closely by age as the second weakest relationship. The second strongest relationship with class thus, price.

Part 6: Reliability of the service and products

6.1 Lafrideradora

Problem 6

Lafrideradora is a subsidiary of the company, who makes components for the company's refrigerators. The blueprint specification for the thickness of a refrigerator part is 0.06 ± 0.04 (cm). It costs $45 to scrap a part if it is outside of the specifications.

The Taguchi loss function was determined:

$L(x) = k(x-T)^2$
$45 = k(0.04)^2$
$k = 28125$
$L(x) = 28125(x-T)^2$

This Taguchi loss function follows a quadratic equation and indicates as the specification moves away from the 0.06cm it will become more expensive and is scrapped at 0.06±0.04cm at cost of $45.

Problem 7

After finding a way to reduce the scrap cost $35

The Taguchi loss function was determined:

$L(x) = k(x-T)^2$
$35 = k(0.04)^2$
$k = 21875$
$L(x) = 21875(x-T)^2$

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

In a) above the Taguchi loss function follows a quadratic equation and indicates as the specification moves away from the 0.06cm it will become more expensive and is scrapped at 0.06±0.04cm at cost of $35. From b) above it indicates that when reducing the deviation target for the thickness specification from 0.04cm to 0.027cm, it will result in a lower scrap cost of $15.95 compared to $35. Thus it will be advantageous for the company to implement these changes.

6.2 Magnaplex

Another subsidiary of the company, Magnaplex, manufactures some parts for the company's technology items.

Reliability of machine A, Ra = 0.85
Reliability of machine B, Rb = 0.92
Reliability of machine C, Rc = 0.90

System reliability if only one machine at each stage is operational:
R(s) = Ra*Rb*Rc = 0.7038

System reliability when having two machines at each stage
$R(s) = (1-(1-Ra)^2)*(1-(1-Rb)^2)*(1-(1-Rc)^2) = 0.9615$

Thus, it is clear that the system is much more reliable when there are two machines in parallel at each stage.

6.3 For delivery process

For the delivery process, there are 20 delivery vehicles available, of which 19 is required to be operating at any time to give a 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.
There are also 21 drivers, who each work an 8 hour shift per day. During the past 1560 days, the number of days that there were only 20 drivers available was 95 days, only 19 drivers available was
6 days and only 18 drivers available, once only.
Estimate on how many days per year we should expect reliable delivery times, given the information above. If we increased our number of vehicles by one to 21, how many days per year we should expect reliable delivery times?

Given the information above, the binomial distribution needs to be used to solve for the number of days per year we should expect reliable delivery times

First we will determine the number of days that vehicles will be reliable independent of the drivers and then we will determine the number of days that drivers will be reliable independent of the vehicles. Then we will combine the two respective reliabilities to determine the overall reliability of deliveries and be able to estimate the number od days we can expect reliable delivery times.

Vehicles:
Vehicles will be reliable if there are 0 or 1 failures.
The vehicle reliability probability = 0.007423949

As a result the number of days per year where there are no failures = 314.4624 and the number of days per year where there is one failure = 47.04028. Thus the number of days per year where the delivery vehicles will be reliable = 361.5027

Drivers:

Drivers will be reliable if there are 0, 1 or 2 failures

The driver reliability probability = 0.003224402

As a result the number of days per year where there are no failures = 341.0658, the number of days per year where there is one failure = 23.16911 and the number of days per year where there are two failures = 0.7494819. Thus the number of days per year where the delivery drivers are reliable = 364.9844

Overall reliability:

The overall delivery reliability = total vehicle reliability * total driver reliability

= 0.9904183 * 0.9999573

= 0.990376

Thus, the total number of days per year we can expect reliable delivery times = 361.4872

This means that in a year of 365 days the delivery times will be reliable for 361 of those days.

If we increased the number of vehicles to 21 but the vehicle failure probability stays the same:

Vehicles will be reliable if there are 0, 1 or 2 failures
The vehicle reliability probability = 0.007423949

As a result the number of days per year where there are no failures = 312.1278, the number of days per year where there is one failure = 49.02561 and the number of days per year where there are two failures = 3.666858. Thus the number of days per year where the delivery vehicles are reliable = 364.8203

New overall reliability with 21 vehicles:

The overall delivery reliability = total new vehicle reliability * total driver reliability

= 0.9995077 * 0.9999573

= 0.999465

Thus, the total number of days per year we can expect reliable delivery times = 364.8047

This means with one additional vehicle, in a year of 365 days the delivery times will be reliable for 364 of those days.

Conclusion

After completing the data wrangling and a descriptive analysis of the valid data set, it was discovered that household and gift items had an increasing trend in delivery times which should be looked into by management as this is concerning. Luxury items, which brings in the most revenue, had a decreasing trend in delivery times which shows a continuous delivery time process improvement or human error when capturing data.

In addition, the technology items delivery time process is currently not capable of meeting the required delivery specifications with a CpK value of less than one that management should seek to improve.

The delivery times for technology was optimized to maximise profits and it can be concluded that the average delivery time needs to be reduced by two days in order to minimize penalty costs for late deliveries.

Furthermore a MANOVA test was done on all the things vs the class of item and it was found that the days and months made no significant difference while the age, year, delivery time and price did make a significant effect on the class of item bought.

Two subsidiaries of the company were evaluated and recommendations were made to improve their respective targets and methods. Finally the reliability of the delivery times were calculated with respect to the vehicles and drivers independently as well as in combination providing an overall reliability, it was found that reliable delivery can be expected for 361 days of the year. However, if an additional vehicle is added the reliability increases to 364 days per year.

References

Evan, J. (2017). *Managing for Quality and Performance Excellence 10 Ed.* Cengage Learning, Inc.

Hernandez, F. (2015, January 10). *Data analysis with R - exercises*. Retrieved from http://fch808.github.io/Data-Analysis-with-R-Exercises.html

MANOVA Test in R: Multivariate Analysis of Variance. (n.d.). Retrieved from STHDA: http://www.sthda.com/english/wiki/manova-test-in-r-multivariate-analysis-of-variance