UNIVERSITEIT·STELLENBOSCH·UNIVERSITY
jou kennisvennoot • your knowledge partner

Quality Assurance 344

ECSA Project

Authors
Walter R. Nel

23580607

19 October 2022

Plagiarism declaration

I have read and understand the Stellenbosch University Policy on Plagiarism and the definitions of plagiarism and self-plagiarism contained in the Policy [Plagiarism: The use of the ideas or material of others without acknowledgement, or the re-use of one's own previously evaluated or published material without acknowledgement or indication thereof (self-plagiarism or text-recycling)].

I also understand that direct translations are plagiarism, unless accompanied by an appropriate acknowledgement of the source. I also know that verbatim copy that has not been explicitly indicated as such, is plagiarism.

I know that plagiarism is a punishable offence and may be referred to the University's Central Disciplinary Committee (CDC) who has the authority to expel me for such an offence.

I know that plagiarism is harmful for the academic environment and that it has a negative impact on any profession.

Accordingly all quotations and contributions from any source whatsoever (including the internet) have been cited fully (acknowledged); further, all verbatim copies have been expressly indicated as such (e.g. through quotation marks) and the sources are cited fully.

I declare that, except where a source has been cited, the work contained in this assignment is my own work and that I have not previously (in its entirety or in part) submitted it for grading in this module/assignment or another module/assignment.

I declare that have not allowed, and will not allow, anyone to use my work (in paper, graphics, electronic, verbal or any other format) with the intention of passing it off as his/her own work.

I know that a mark of zero may be awarded to assignments with plagiarism and also that no opportunity be given to submit an improved assignment.

Signature:

Name: Walter Riaan Nel .... Student no: 23580607

Date: 19/10/2022

Abstract

This report demonstrates and explains the data analytics techniques used to evaluate a dataset that contains client data for an online business.

The report shows the 6 different parts of the project undertaken which includes data wrangling, descriptive statistics, statistical process control, optimizing delivery processes, DOE and MANOVA tests and reliability of services and products.

Table of Contents

List of figures

List of Tables

1. Introduction

One of the main objectives for an industrial engineer is to optimize processes. It is important for industrial engineers to obtain the skills of analyzing data in order to optimize the associated systems involved in producing the data. This report will demonstrate the processes involved in analyzing client data for an online business. It will describe the processes and outcomes of data wrangling where instances with missing values are removed, descriptive statistics where trends and patterns in the data are recognized, statistical process control, optimizing delivery processes, DOE and MANOVA tests and reliability of services and products. The report will end in a conclusion followed by references and an R appendix.

1. Part 1: Data Wrangling

Given the data of *QuallitySales2022,* the data was wrangled to remove all missing values from the dataset. Therefore, the data was split into Valid Data (instances that do not contain missing values) and Invalid Data (instances that contain missing values).

1.1 Valid Data

NewIndex	X	ID	AGE	Class	Price	Year	Month	Day	elivery.tim	Why.Bought
1	1	19966	54	Sweets	246.21	2021	7	3	1.5	Recommended
2	2	34006	36	Household	1708.21	2026	4	1	58.5	Website
3	3	62566	41	Gifts	4050.53	2027	8	10	15.5	Recommended
4	4	70731	48	Technology	41843.21	2029	10	22	27	Recommended
5	5	92178	76	Household	19215.01	2027	11	26	61.5	Recommended
6	6	50586	78	Gifts	4929.82	2027	4	24	14.5	Random
7	7	73419	35	Luxury	108953.5	2029	11	13	4	Recommended
8	8	32624	58	Sweets	389.62	2025	7	2	2	Recommended
9	9	51401	82	Gifts	3312.11	2025	12	18	12	Recommended
10	10	96430	24	Sweets	176.52	2027	11	4	3	Recommended
11	11	87530	33	Technology	8515.63	2026	7	15	21	Browsing
12	12	14607	64	Gifts	3538.66	2026	5	13	13.5	Recommended
13	13	24299	52	Technology	27641.97	2024	5	29	17	Browsing
14	14	77795	92	Food	556.83	2025	6	3	3	Random
15	15	62567	73	Clothing	347.99	2024	3	29	8.5	Website
16	16	14839	47	Technology	54650.41	2027	12	30	18.5	Recommended
17	17	96208	44	Technology	14739.09	2028	3	17	13	Recommended
18	18	39674	69	Technology	22315.17	2026	8	20	20.5	Recommended
19	19	98694	74	Sweets	546.48	2025	5	9	2	Recommended
20	20	99187	54	Luxury	81620.21	2027	9	14	3	Recommended
21	21	59365	72	Gifts	3314.76	2028	4	30	13	Recommended
22	22	37221	24	Sweets	220.91	2021	3	8	3	Recommended
23	23	78120	23	Gifts	2378.31	2023	3	10	12	Recommended
24	24	65860	30	Gifts	2440.41	2021	5	11	9.5	Recommended
25	25	70953	70	Gifts	3962.67	2024	10	6	12.5	Recommended

Table 1: Valid Data

The valid data consists of 179983 instances. A new column was added to indicate the new indexing of the valid data. The second column indicates the original indexing of the instances.

1.2 Index Difference

12344	12344	90260	34	Luxury	42891.66	2025	8	4	4	Recommended
12345	12346	92286	32	Technolog	38167.24	2028	7	6	19.5	Website
12346	12347	89263	44	Clothing	891.71	2021	7	2	8.5	Recommended
12347	12348	71191	49	Household	14936.31	2025	10	11	43.5	Recommended
12348	12349	24801	28	Food	425.96	2022	1	29	2.5	Recommended
12349	12350	85475	57	Luxury	78817.55	2026	3	21	5	Browsing
12350	12351	61842	24	Clothing	1008.78	2025	7	16	8	Recommended
12351	12352	49373	34	Technolog	17277.26	2024	10	11	14.5	Browsing
12352	12353	40283	45	Technolog	16930.76	2025	3	9	27.5	EMail
12353	12354	19084	56	Sweets	171.81	2026	10	8	1.5	Random
12354	12355	53251	30	Clothing	322.12	2021	8	26	10	Recommended
12355	12356	21484	63	Gifts	2099.09	2027	1	3	12	Browsing

Table 2: Index Difference

Table 2 indicates the first instance's index to change as instance 12345 is the first instance which contains a missing value.

1.3 Invalid Data

NewIndex	X	ID	AGE	Class	Price	Year	Month	Day	elivery.tim	Vhy.Bought
1	12345	18973	93	Gifts		2026	6	11	15.5	Website
2	16321	81959	43	Technology		2029	9	6	22	Recommended
3	19541	71169	42	Technology		2025	1	19	20.5	Recommended
4	19999	67228	89	Gifts		2026	2	4	15	Recommended
5	23456	88622	71	Food		2027	4	18	2.5	Random
6	34567	18748	48	Clothing		2021	4	9	8	Recommended
7	45678	89095	65	Sweets		2029	11	6	2	Recommended
8	54321	62209	34	Clothing		2021	3	24	9.5	Recommended
9	56789	63849	51	Gifts		2024	5	3	10.5	Website
10	65432	51904	31	Gifts		2027	7	24	14.5	Recommended
11	76543	79732	71	Food		2028	9	24	2.5	Recommended
12	87654	40983	33	Food		2024	8	27	2	Recommended
13	98765	64288	25	Clothing		2021	1	24	8.5	Browsing
14	144444	70761	70	Food		2027	9	28	2.5	Recommended
15	155555	33583	56	Gifts		2022	12	9	10	Recommended
16	166666	60188	37	Technology		2024	10	9	21.5	Website
17	177777	68698	30	Food		2023	8	14	2.5	Recommended

Table 3: Invalid Data

The dataset consists of 17 instances that contains missing values. A new column was added to indicate the new indexing of the invalid data. The second column indicates the original indexing of the instances.

2 Part 2: Descriptive Statistics

The Valid Dataset was analyzed by using descriptive statistics.

2.1 Graphs

Sales Count for each year

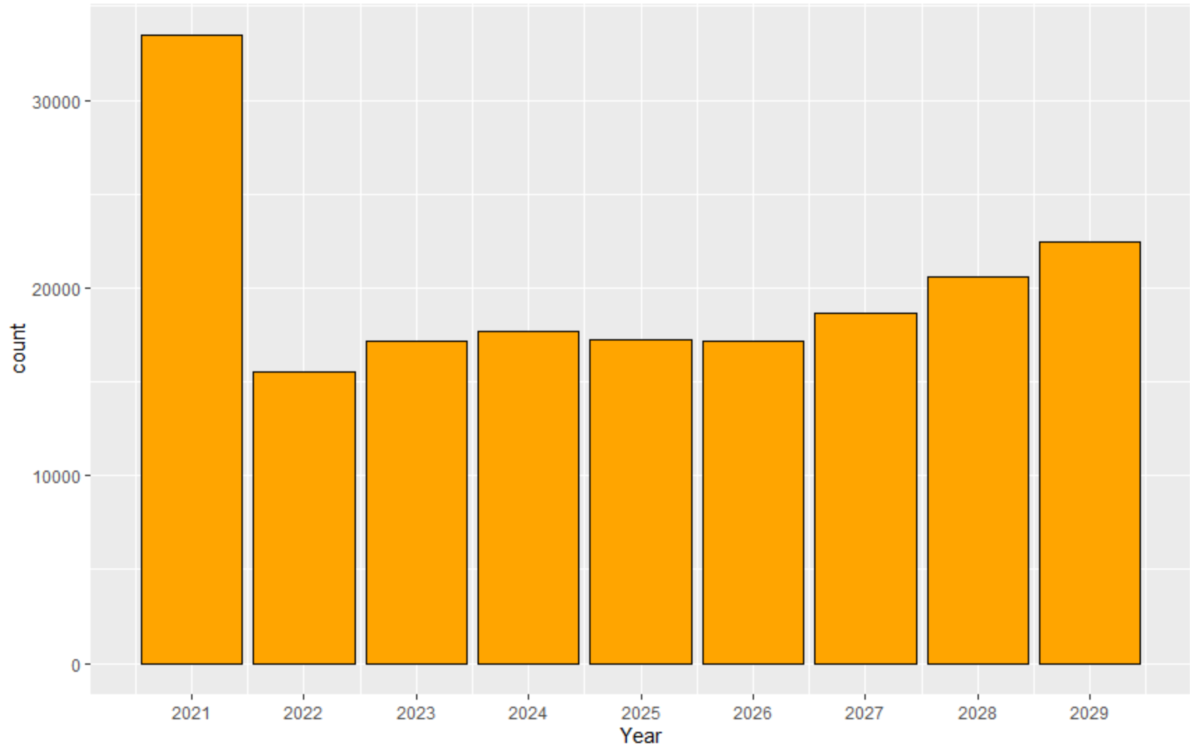


Figure 1: Sales Count vs Year

Figure 2.1 indicates that the most sales were made in the year 2021. After that, the sales increased linearly with a positive trend from 2022 to 2029

Sales Count for each Month

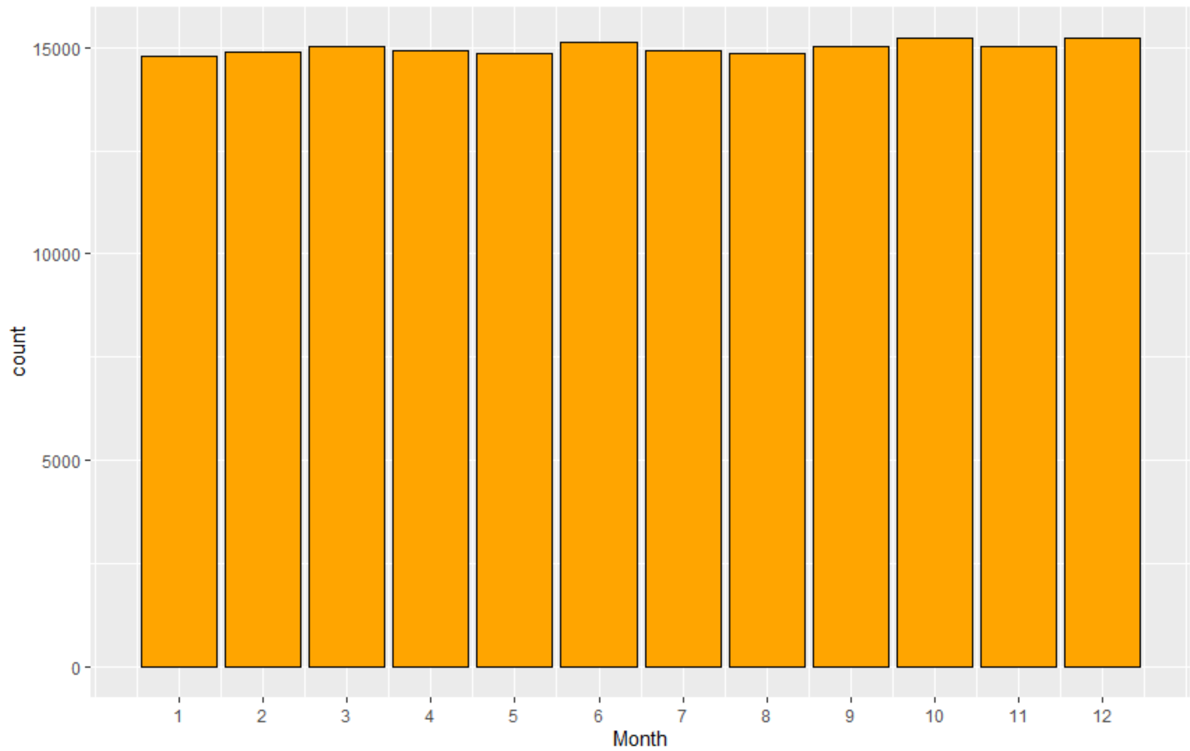


Figure 2: Sales Count vs Month

The Sales Count for each Month is uniformly distributed with no trend

Sales Count for each day of the month

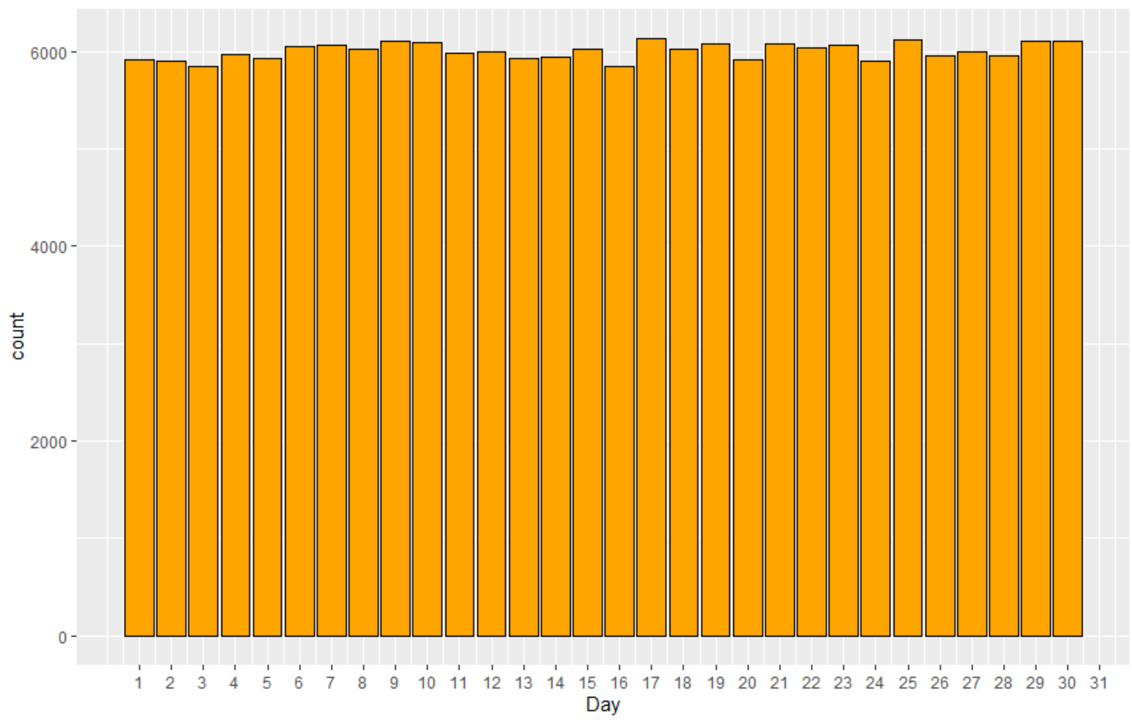


Figure 3: Sales Count vs Day

The Sales Count for each day is uniformly distributed with no recognizable trend.

Sales Count vs Age

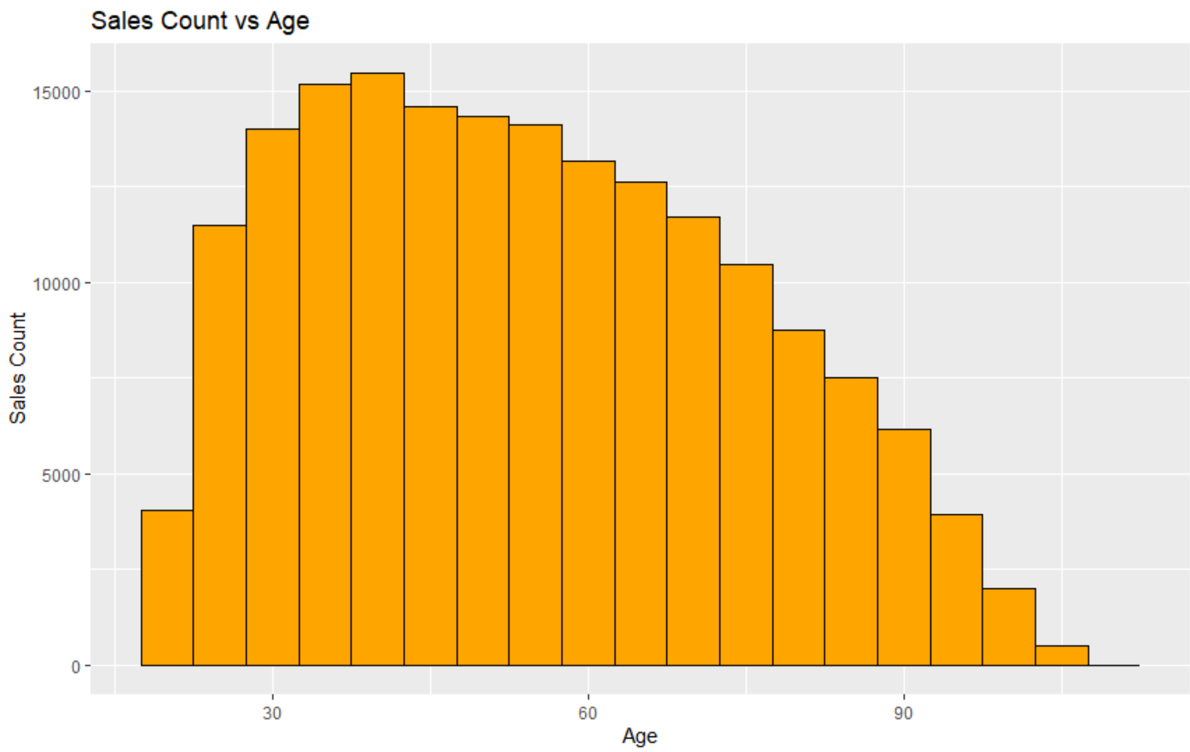


Figure 4: Sales Count vs Age

The distribution of Sales Count vs Age is positively skewed (skewed to the right). This indicates that on average, people are less likely to buy a product as their age increases from 40 years old.

Delivery Time vs Sales

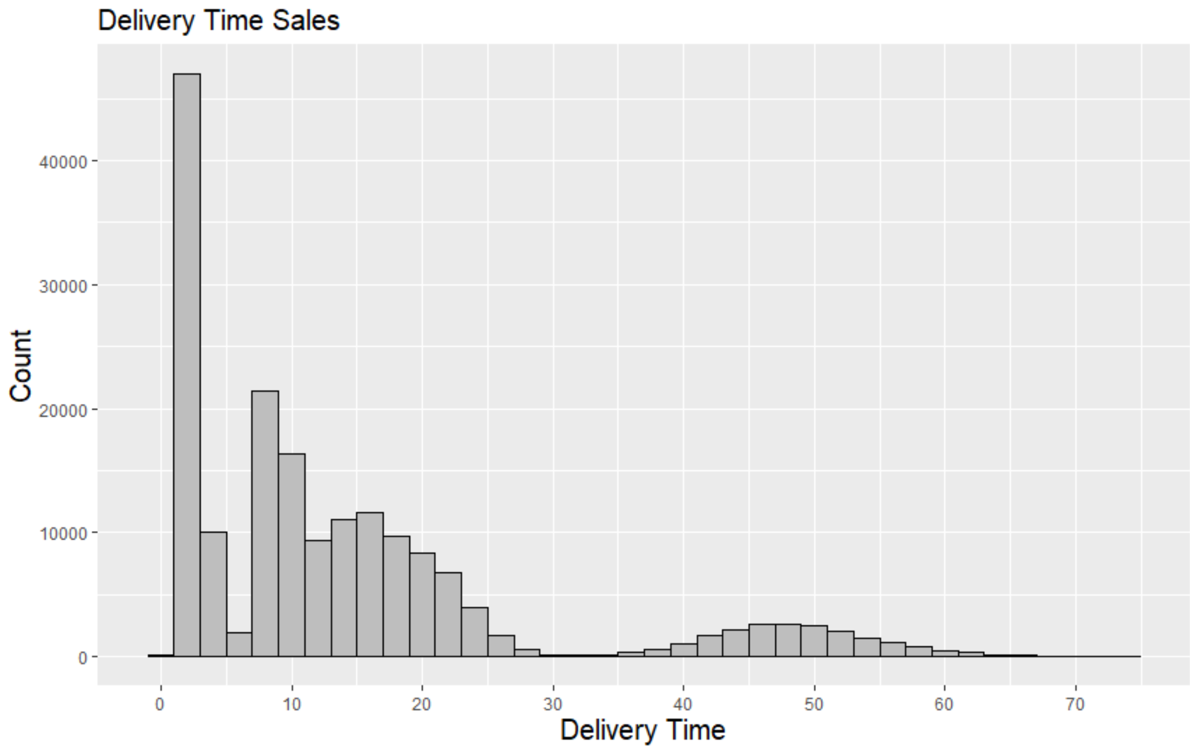


Figure 5: Sales Count vs Delivery Time

The Delivery Time vs Sales has a slight normal distribution between 30 and 70 days.

Reasons for buying

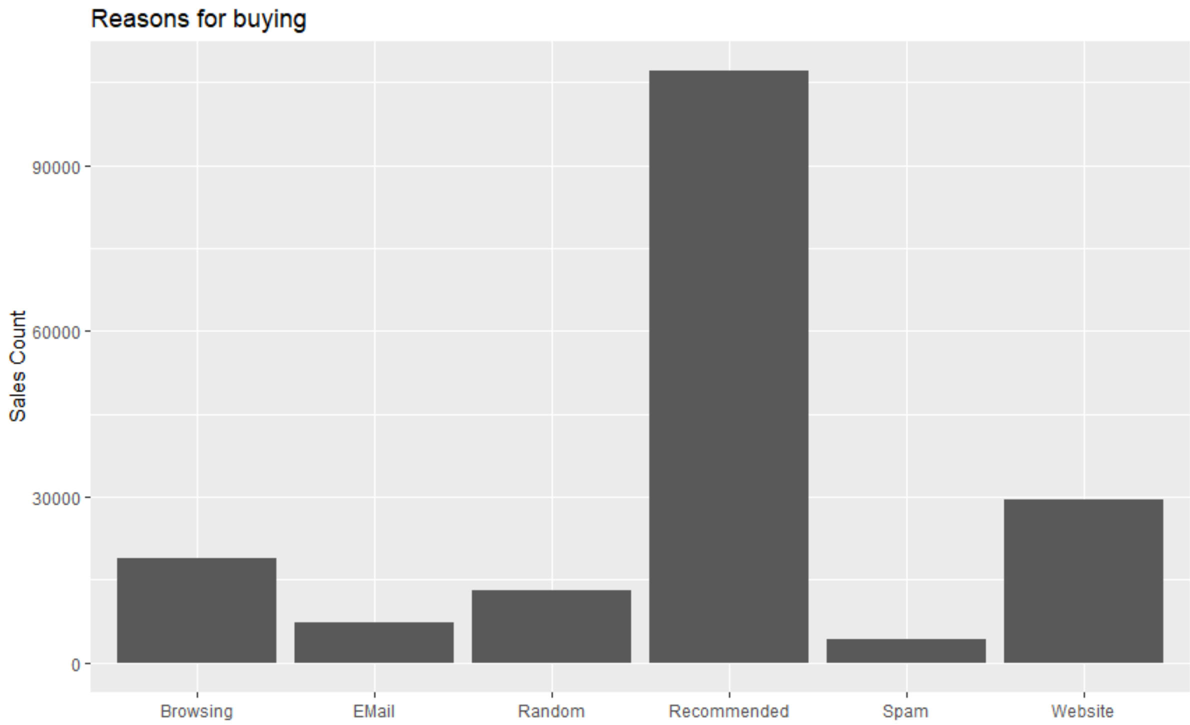


Figure 6: Sales Count vs Reasons for buying

Figure 2.6 indicates that the number one reason for people buying products is that it is recommended.

Sales Count vs Class

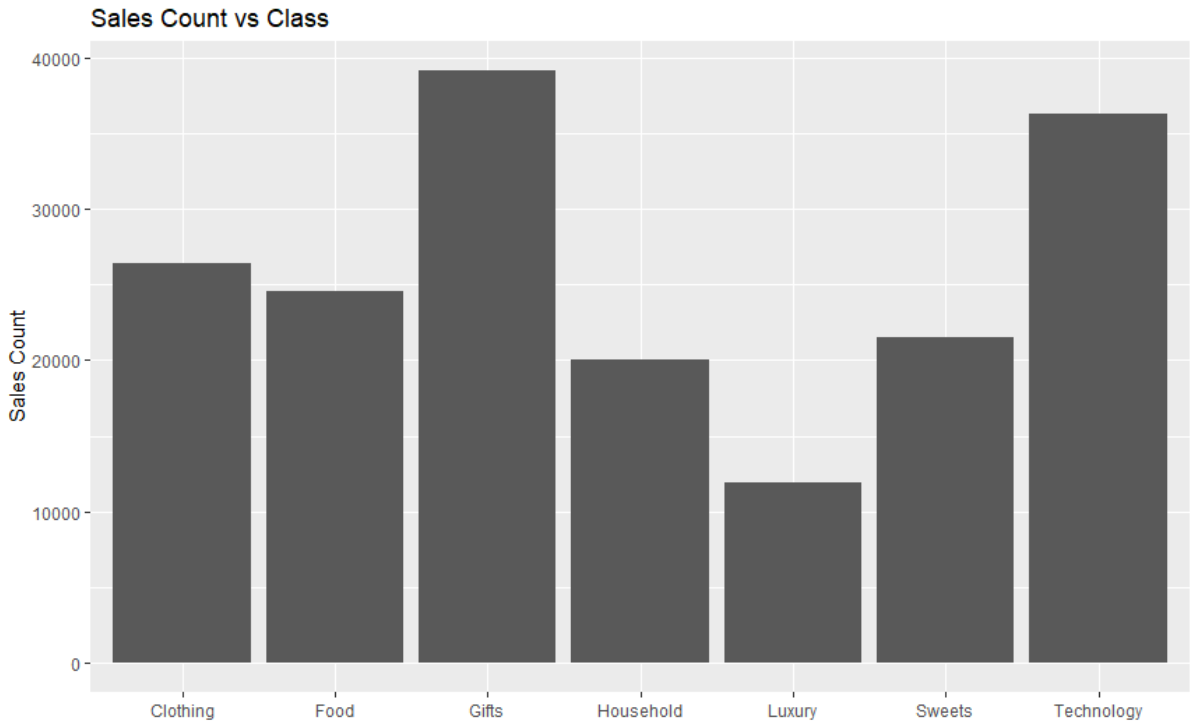


Figure 7: Sales Count vs Class

Figure 2.7 indicates that classes Gifts and Technology has the most sales and that luxury items has the least number of sales.

Age vs Why Bought

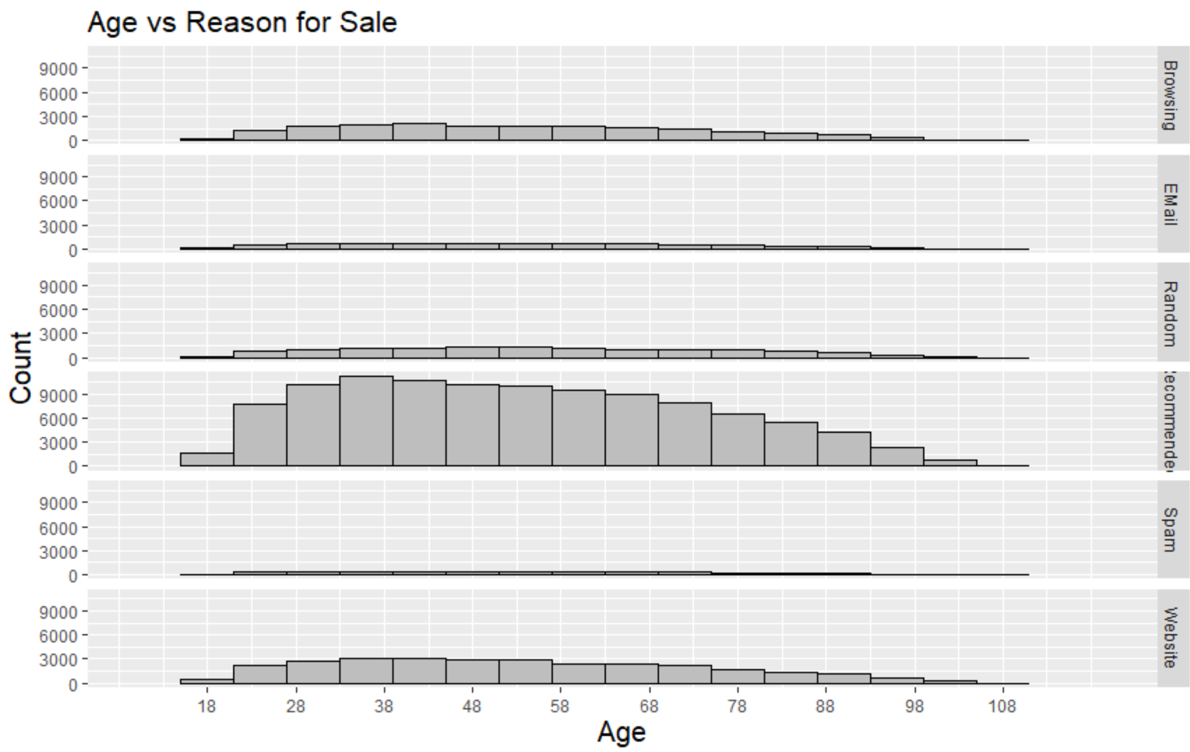


Figure 8: Age vs Why bought

The distribution for each reason of buying a product vs age is positively skewed (skewed to the right). Thus, on average, for each reason for buying a product, the number of sales decreases as the age of people increases.

Age vs Class

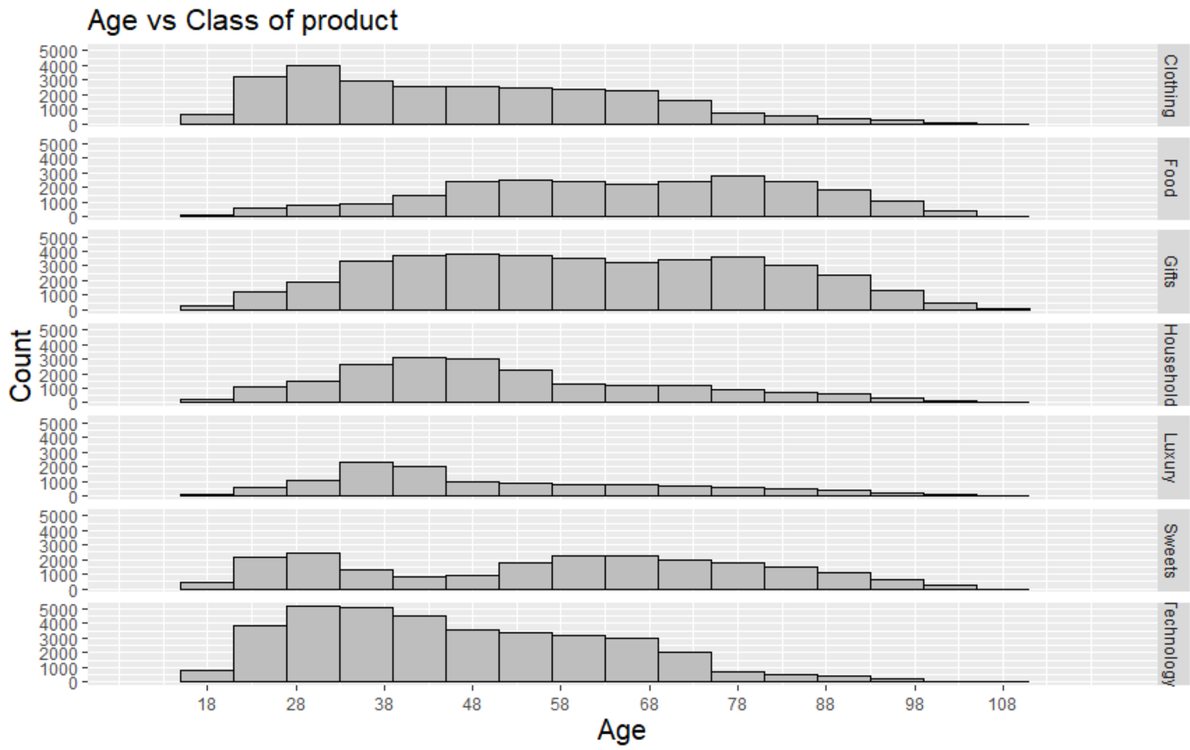


Figure 9: Age vs Class of Productt

A general conclusion is made that younger people (between the ages of 20 and 35) tend to buy more clothing, household, luxury and technology products.

Food, Gifts and sweets are normally distributed between the age groups.

Price vs Class of Product

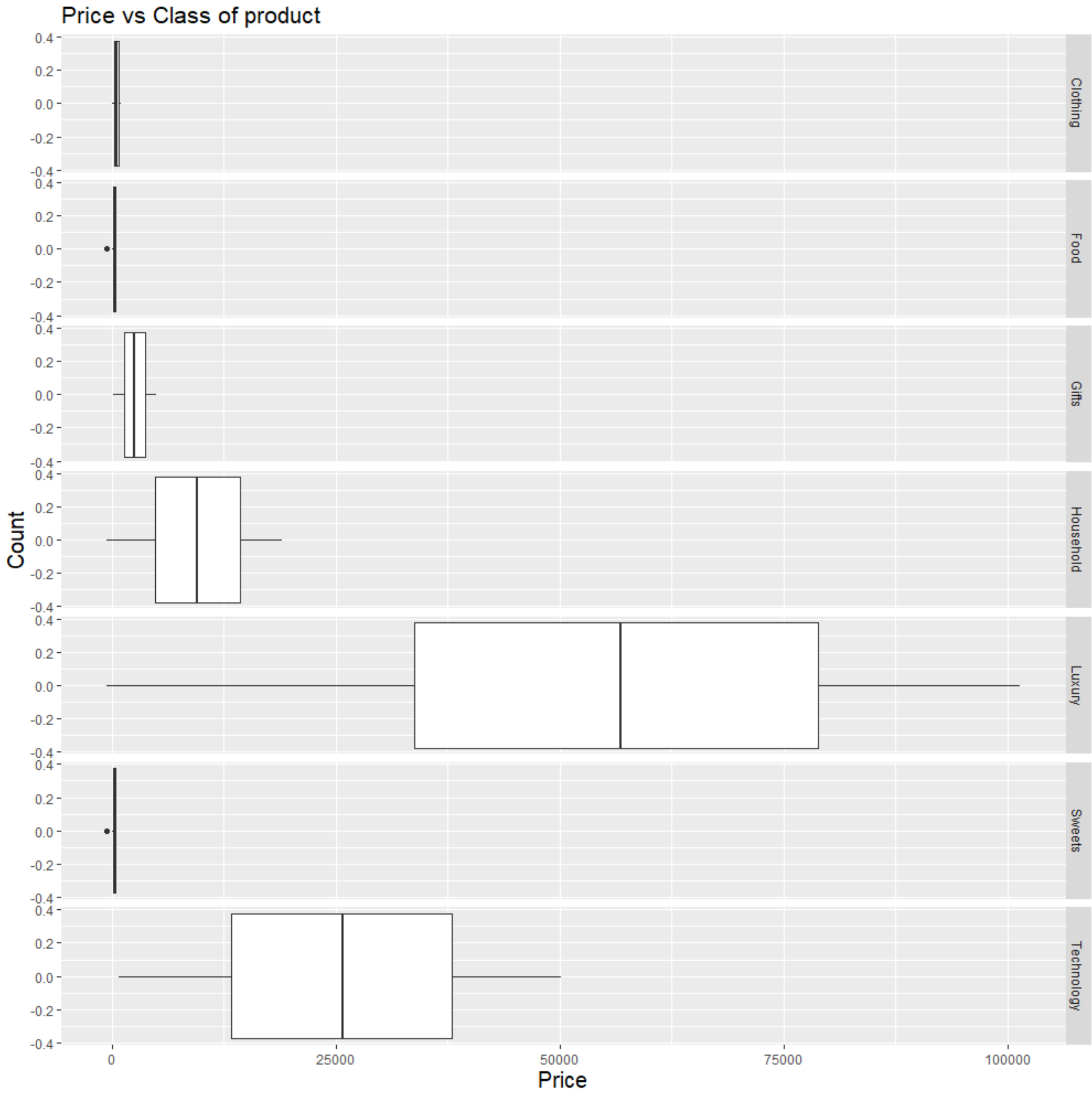


Figure 10: Price vs Class of Product

By analysing figure 2.10, a general conclusion is made that clothing, food and sweets are the classes with the cheapest items.

Luxury items are by far the class with the most expensive items, with technology having the second most expensive items.

Price vs Why Bought

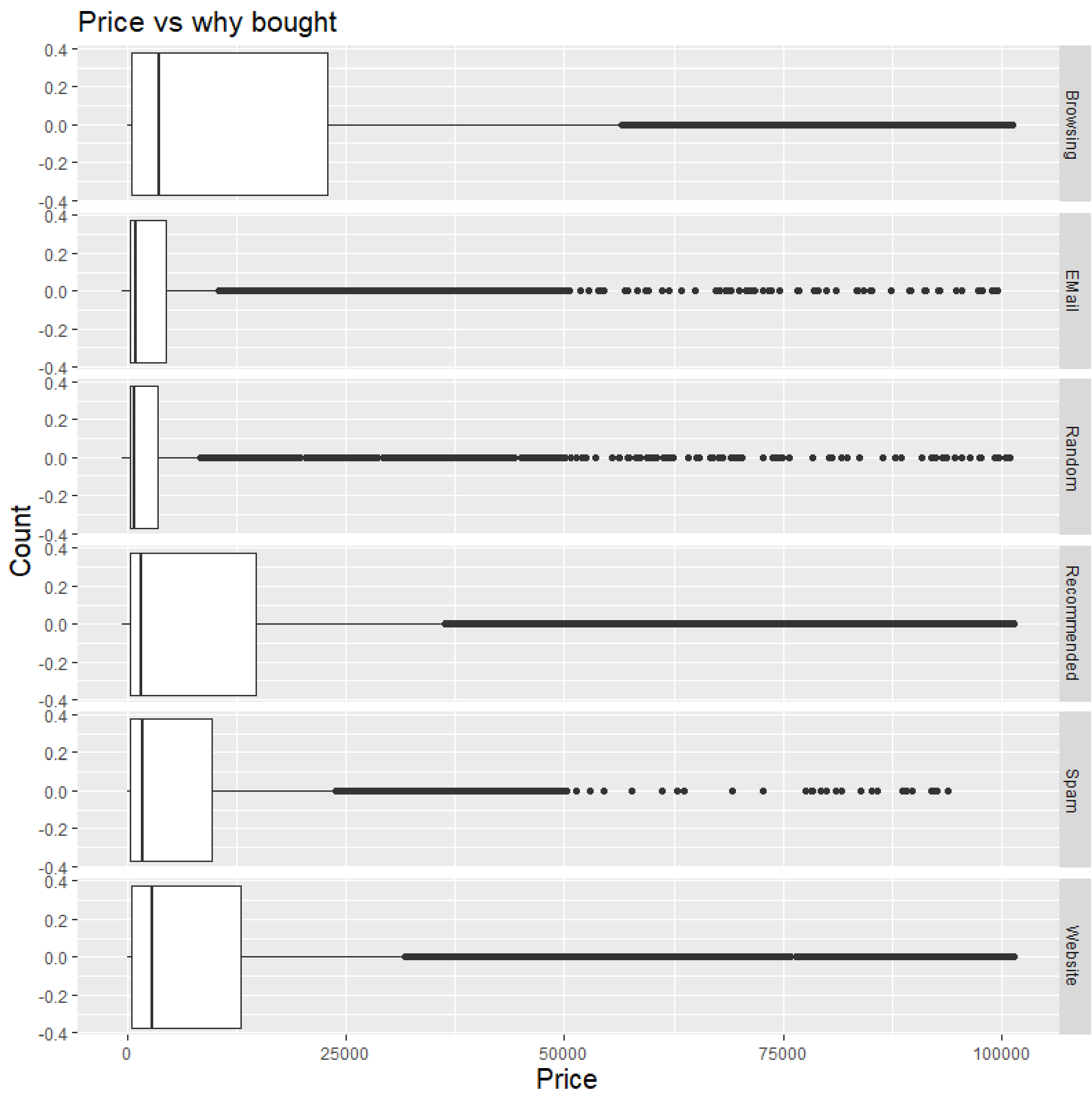


Figure 11: Price vs Why Bought

The box plots overlap each other.

Thus, no specific trend between the prices of products and the reason why it's bought can be identified.

2.2 Process Capability Indices

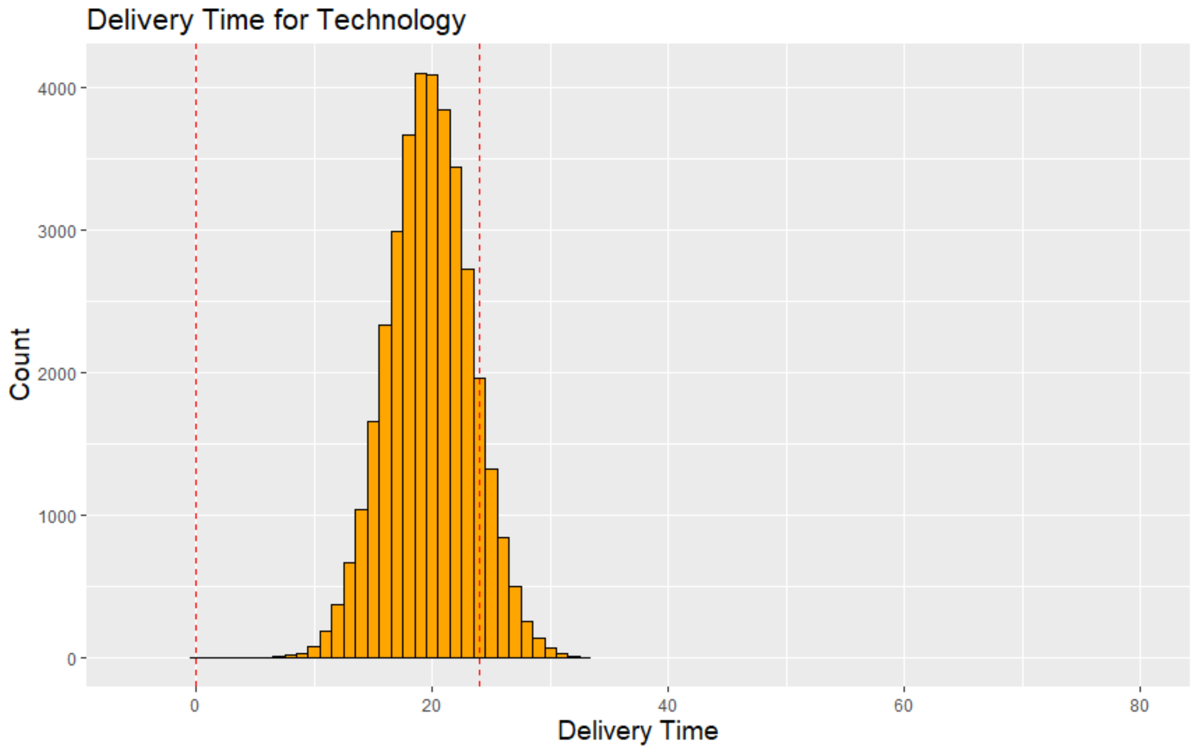


Figure 12: Delivery Time for Technology

Given USL = 24 and LSL = 0,

The following Process Capabilities indices are calculated:

Logically, LSL cannot be less than zero, because time cannot be a negative value.

$Cp = 1.142207$

$Cpu = 0.3796933$

$Cpl = 1.90472$

$Cpk = 0.3796933$

3 Part 3 – Statistical Process Control

In Part 3, statistical process control (SPC) was done to construct X&S-charts.

For delivery process times, the first 30 samples of the dataset were used to determine upper control limits, 2-sigma control limits, 1-sigma control limits, center lines, and lower control limits.

X-Chart

Class	UCL	U2Sigma	U1Sigma	CL	L1Sigma	L2Sigma	LCL
Technology	22.9731	20.80755	21.24066	20.37444	19.50822	19.94133	17.77579
Clothing	9.404681	9.042447	9.114894	8.97	8.825106	8.897553	8.535319
Household	50.24618	47.17621	47.79021	46.56222	45.33423	45.94823	42.87826
Luxury	5.493524	4.861884	4.988212	4.735556	4.4829	4.609228	3.977587
Food	2.70933	2.526555	2.56311	2.49	2.41689	2.453445	2.27067
Gifts	9.487909	8.548911	8.73671	8.361111	7.985512	8.173311	7.234313
Sweets	2.896798	2.547615	2.617451	2.477778	2.338105	2.407941	2.058758

S-Chart

Class	UCL	U2Sigma	U1Sigma	CL	L1Sigma	L2Sigma	LCL
Technology	5.179912	3.60959	3.92366	3.295528	2.98146	2.6674	1.411143
Clothing	0.86645	0.60378	0.65631	0.551247	0.49871	0.44618	0.236044
Household	7.343248	5.1171	5.56233	4.67187	4.22664	3.78141	2.000493
Luxury	1.51086	1.05283	1.14444	0.961229	0.86962	0.77802	0.411598
Food	0.437191	0.30465	0.33116	0.278147	0.25164	0.22513	0.119102
Gifts	2.246048	1.56515	1.70133	1.428965	1.29278	1.1566	0.611882
Sweets	0.835233	0.58203	0.63267	0.531386	0.48075	0.4301	0.227539

3.1 Graphs for the 30 first samples

Technology

X-Chart

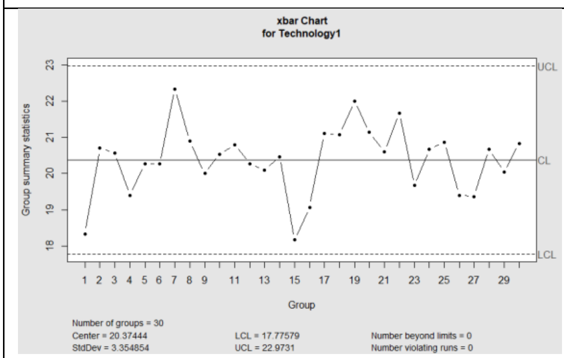
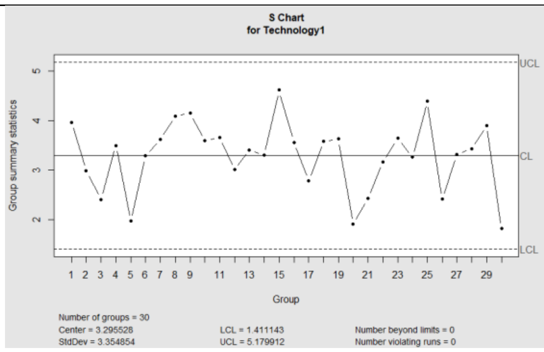


Figure 13: X&S Chart for Technology

S-Chart



Instances are oscillating around the center line of the mean and vriance therefore the X-chart and the S-chart indicates that the first 30 samples of Technology are in control

Clothing

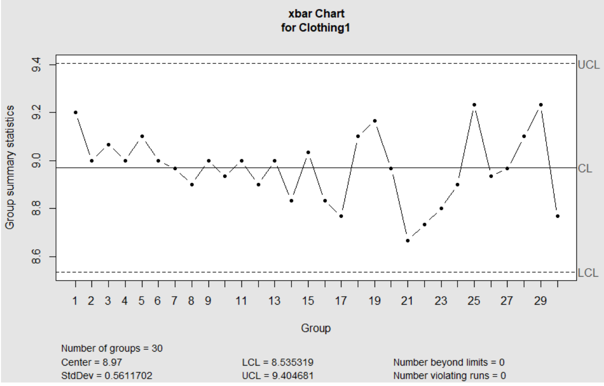
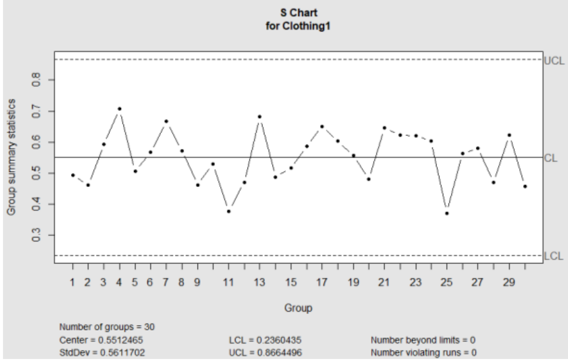
X-Chart	S-Chart
	

Figure 14: X&S Chart for Clothing

Instances are oscillating around the center line of the mean and vriance therefore the X-chart and the S-chart indicates that the first 30 samples of Clothing are in control.

Household

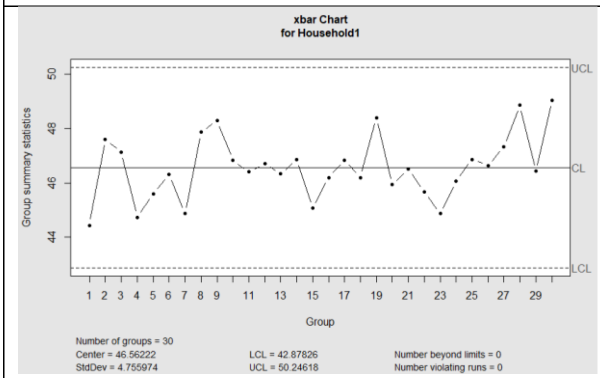
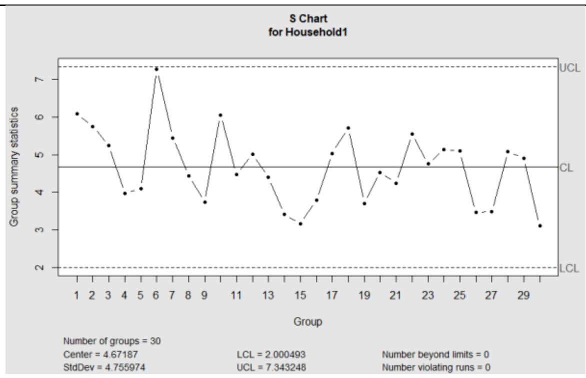
X-Chart	S-Chart
	

Figure 15: X&S Chart for Household

Instances are oscillating around the center line of the mean and vriance therefore the X-chart and the S-chart indicates that the first 30 samples of Household are in control.

Luxury

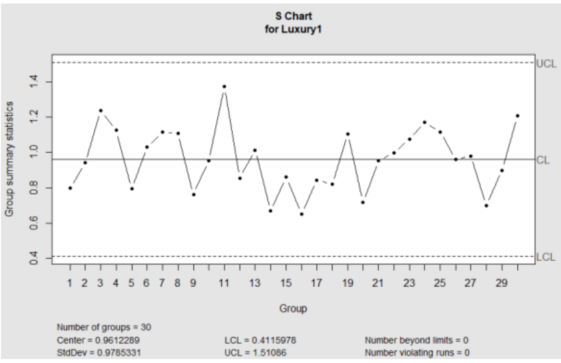
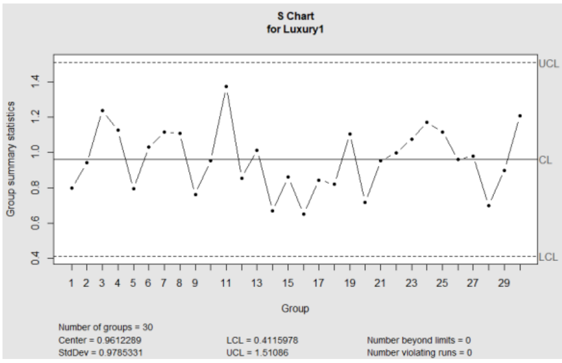
X-Chart	S-Chart
	

Figure 16: X&S Chart for Luxury

Instances are oscillating around the center line of the mean and vriance therefore the X-chart and the S-chart indicates that the first 30 samples of Luxury are in control.

Food

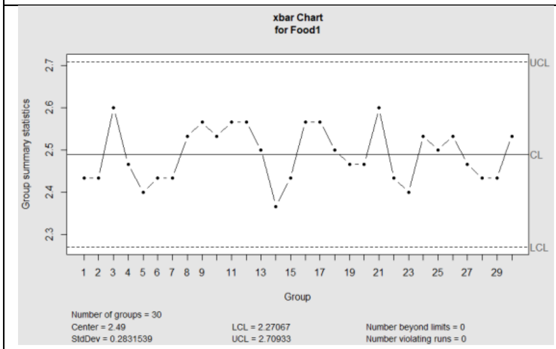
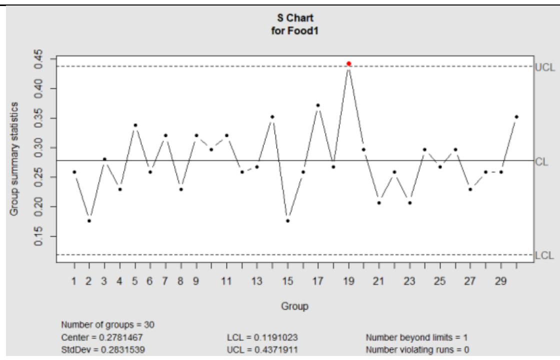
<table>
<tr><th>X-Chart</th><th>S-Chart</th></tr>
<tr><td>


Figure 17: X&S Chart for Food
</td><td>


</td></tr>
<tr><td colspan="2">Instances are oscillating around the center line of the mean and vriance therefore the X-chart and the S-chart indicates that the first 30 samples of Food are in control.</td></tr>
</table>

Gifts

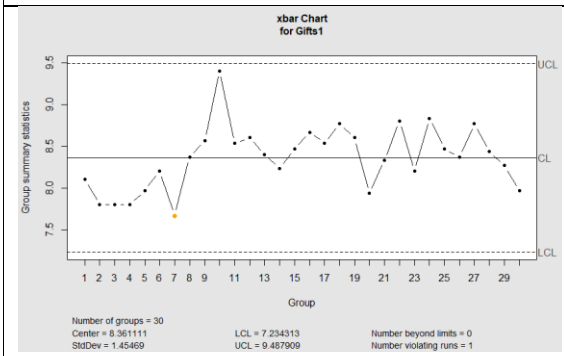
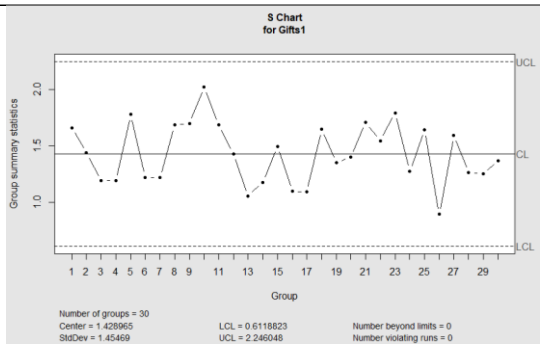
<table>
<tr><th>X-Chart</th><th>S-Chart</th></tr>
<tr><td>


Figure 18: X&S Chart for Gifts
</td><td>


</td></tr>
<tr><td colspan="2">Instances are oscillating around the center line of the mean and vriance therefore the X-chart and the S-chart indicates that the first 30 samples of Gifts are in control.</td></tr>
</table>

Sweets

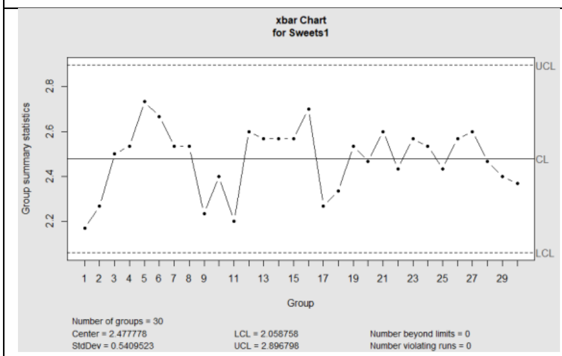
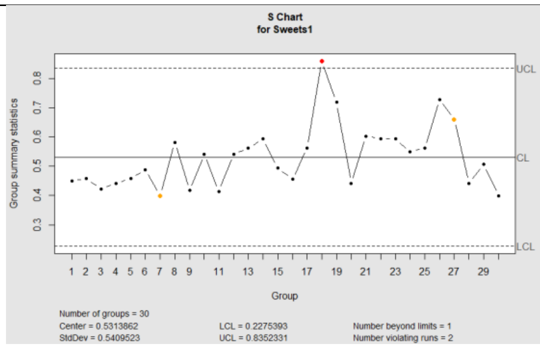
<table>
<tr><th>X-Chart</th><th>S-Chart</th></tr>
<tr><td>


Figure 19: X&S Chart for Sweets
</td><td>


</td></tr>
<tr><td colspan="2">Instances are oscillating around the center line of the mean and vriance therefore the X-chart and the S-chart indicates that the first 30 samples of Sweets are in control.</td></tr>
</table>

3.2 Graphs on all the data

Technology

X-Chart	S-Chart

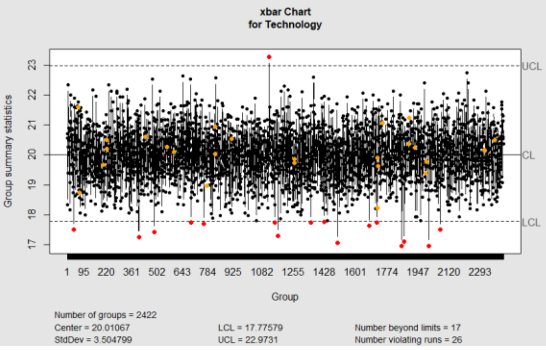
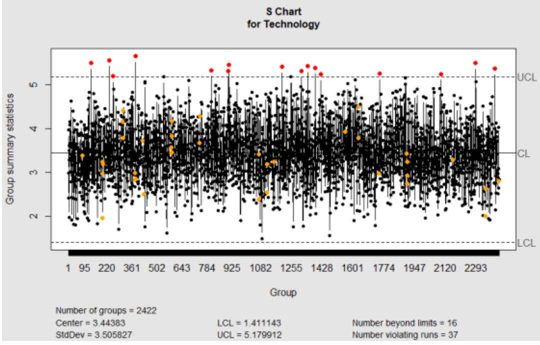


Figure 20: X&S Chart for Technology (All Samples)



The delivery time for Technology shows a stationary series, the mean and variance is not a function of time. Technology is under control with only a few instances outside control limits.

Clothing

X-Chart	S-Chart

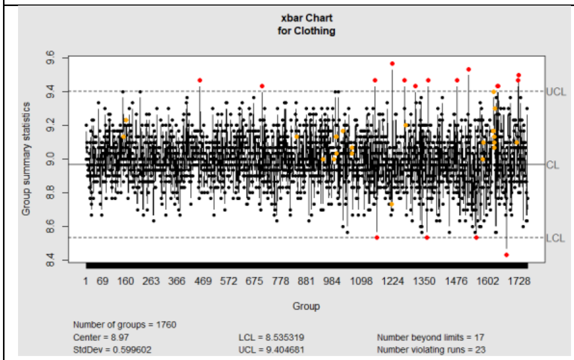
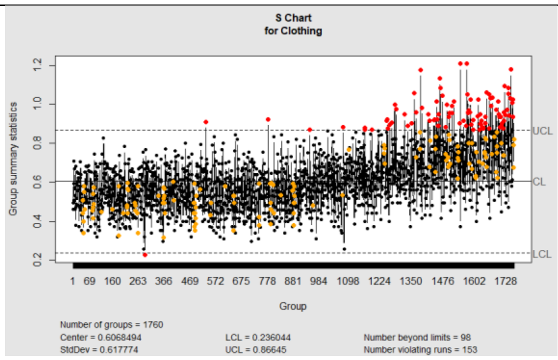


Figure 21: X&S Cart for Clothing (All Samples)



All though the majority of samples are within control limts, the S-chart indicates that clothing is a second order non-stationary time-series, the variance is a function of time with increasing variance at the last instances. This shows that there is more variation towards the end for delivery time of clothing.

Household

X-Chart	S-Chart

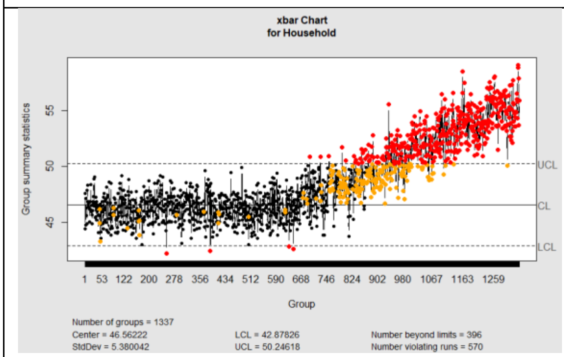
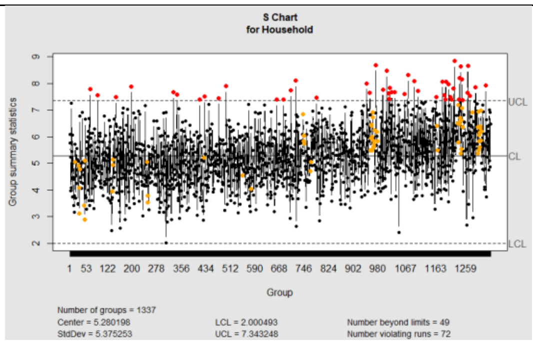


Figure 22: X&S Chart for Household (All Samples)



The delivery time for household items shows an exponential increase in trend for the second half of instances. Investigation is needed to determine why the mean delivery time for household items increased. Thus, delivery time for household products are not in control.

Luxury

X-Chart

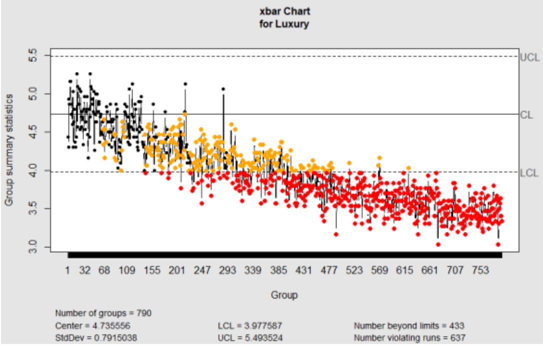
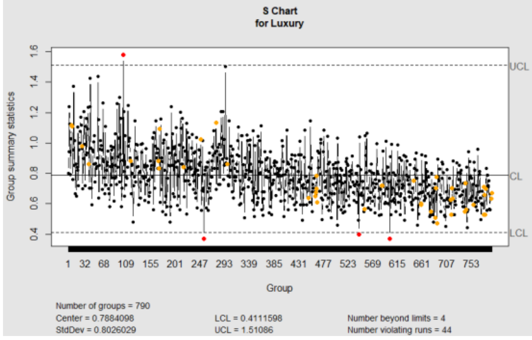


Figure 23: X&S Chart for Luxury (All Samples)

S-Chart



The delivery time for Luxury items shows a linear decrease in trend and is. This could be seen as a positive aspect because a decrease in delivery time for a high valued product would result in a higher return of revenue.

Food

X-Chart

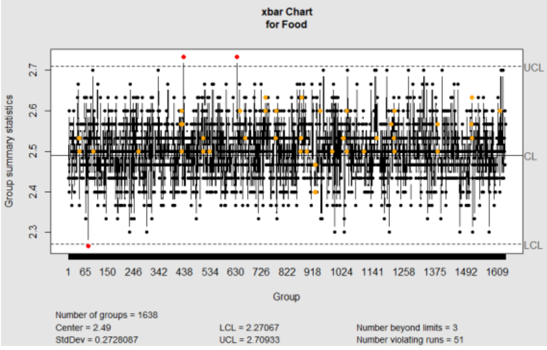
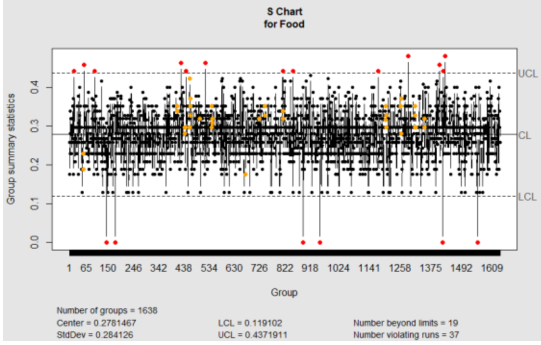


Figure 24: X&S Chart for Food (All Samples)

S-Chart



The delivery time for Food shows a stationary series, the mean and variance is not a function of time. Food is under control with only a few instances outside control limits.

Gifts

X-Chart

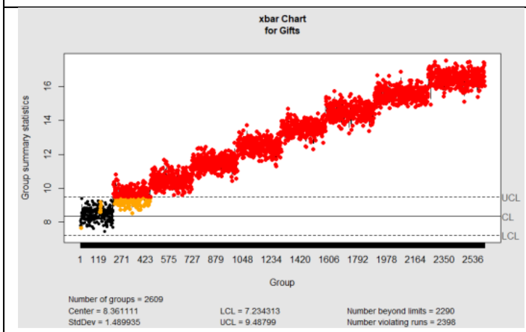
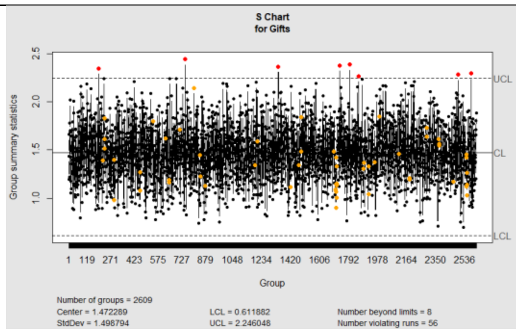


Figure 25: X&S Chart for Gifts (All Samples)

S-Chart



The delivery time for Gifts shows a linear trend and implies first order non-stationary data. This is alarming and further investigation is needed to deterine why the delivery time increases. The charts also indicates that the delivery time for Gifts are not under control.

Sweets

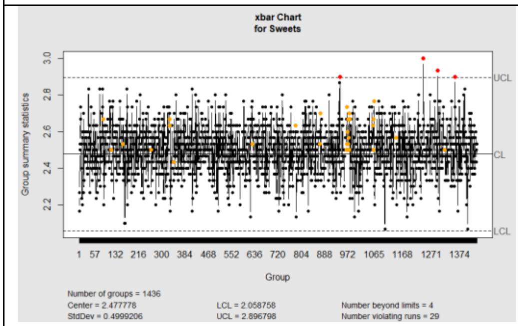
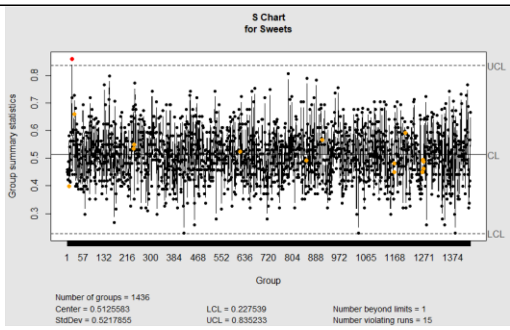
<table>
<tr><th>X-Chart</th><th>S-Chart</th></tr>
<tr><td>

Figure 26: X&S Chart for Sweets (All Samples)</td><td>
</td></tr>
<tr><td colspan="2">The delivery time for Sweets shows a stationary series, the mean and variance is not a function of time. Sweets is under control with only a few instances outside control limits.</td></tr>
</table>

4. Part 4: Optimizing Delivery Processes

4.1 Sample numbers that give indications of out of control

Class	Total	1st Sampl	2nd Sample	3rd Sample	3rd last Sample	2nd last Sample	Last Sample
Technology	17	37	398	483	1872	2009	2071
Clothing	17	455	702	1152	1677	1723	1724
Household	396	252	387	629	1335	1336	1337
Luxury	433	142	171	184	788	789	790
Food	3	75	432	633	NA	NA	NA
Gifts	2290	213	216	218	2607	2608	2609
Sweets	4	942	1243	1294	NA	NA	1358

Table 4: Out of Control Sample Numbers

Household:

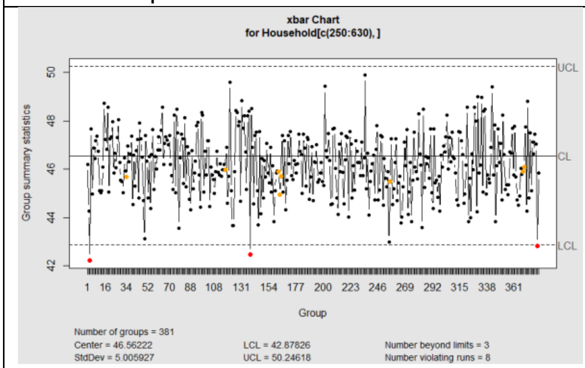
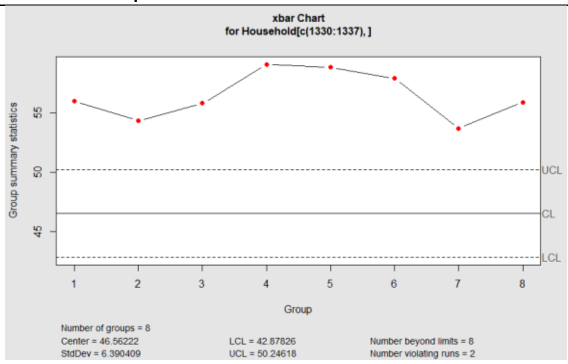
<table>
<tr><th>First 3 Samples</th><th>Last 3 Samples</th></tr>
<tr><td>

Figure 27: First and Last 3 out of Control Limits for Household

Between samples 250 and 630</td><td>

Between samples 1330 and 1337</td></tr>
</table>

Luxury:

First 3 Samples	Last 3 Samples

First 3 Samples

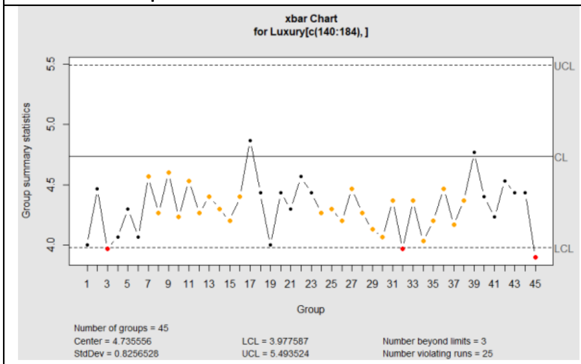
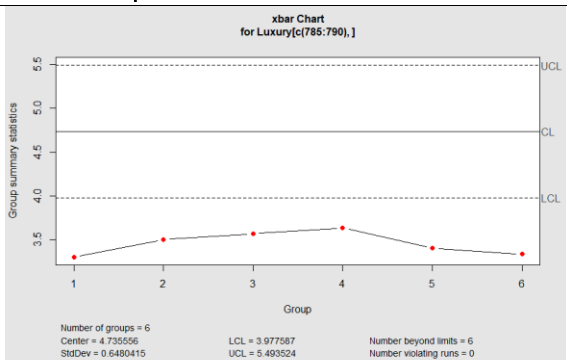


Figure 28: First and Last 3 out of Control Limits for Luxury

Between samples 140 and 184

Last 3 Samples



Between samples 785 and 790

Gifts

First 3 Samples

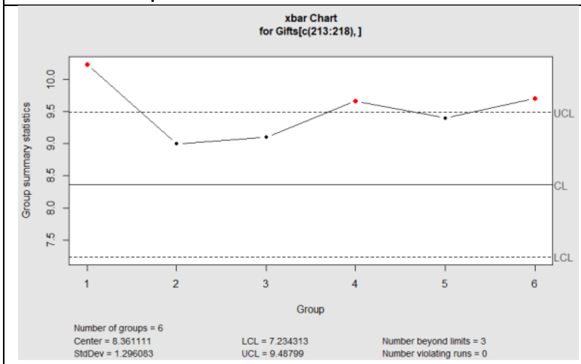
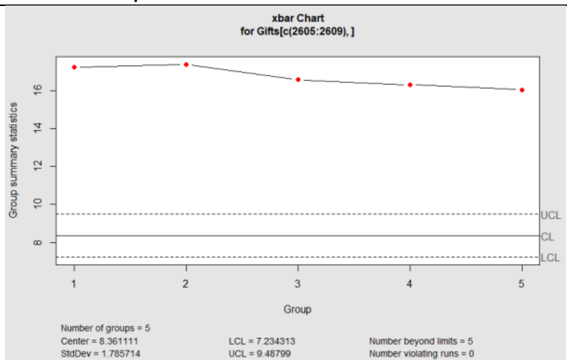


Figure 29: First and Last 3 out of Control Limits for Gifts

Between samples 213 and 218

Last 3 Samples



Between samples 2605 and 2609

4.1. B Most consecutive samples of "sample standard deviations"

Class	Max consecutive	Ending Sample index numver
Technology	6	372
Clothing	4	1013
Household	3	843
Luxury	4	63
Food	6	441
Gifts	5	1651
Sweets	4	971

Table 5: Consecutive "Sample Standard Deviations"

4.2. Likelihood of making type I error for A and B

Null Hypothesis assumed: calculated using the first 30 samples.

- H0: The process is in control and centred on the centreline.
- H1: is that the process in not in control and has moves from the centreline or has increased or decreased in variation.

A type I error (manufacturing error) indicates the probability of stating that the product is not delivered in time when in reality, it is delivered on time.

For A: $0.002699796 = 0.27\%$

For B: $0.7266668 = 72.67\%$

4.3. Minimized delivery cost

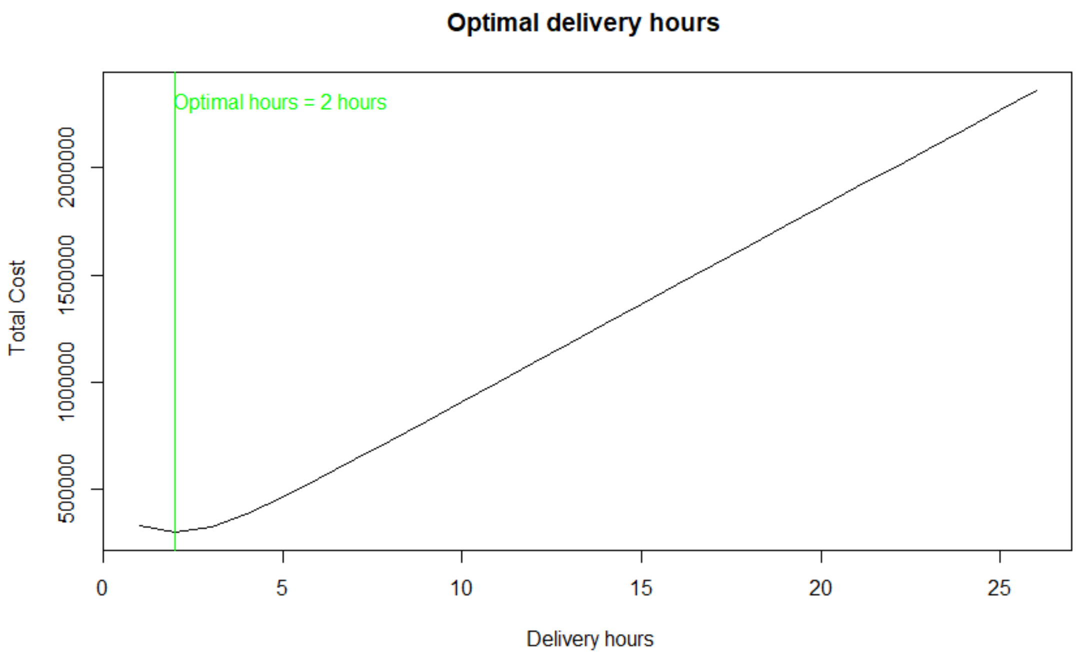


Figure 30: Optimal Delivery Days

By looping through each hour, the costs are compared for different days. The mean number of hours for delivery time is currently 20.01095. The number of sales that are beyond 26 hours is 1356. Thus, 1356 sales are out of control. Each sale that is beyond 26 hours contributes R329 to loss. The current loss is R 446 124 because of loss of sales over 26 hours. It costs R2.5/item/hour to reduce the average time by one hour. To eliminate the loss and move the whole distribution before 26 hours, it will cost R636 072.50

The minimum delivery cost for the technology needs to be calculated by finding the optimal solution between decreasing the loss in Sales (R329/item-late-hour) and taking into effect the moving of the distribution to the left (R2.5/item/hour). This is determined by looping between possible delivery times and estimating the costs associated with it.

Optimal solution: As seen in *figure 30*, the average of the delivery time distribution should be reduced by 2 hours. This will reduce the average delivery time of 20.01095 hours to 18.01095 hours and will minimise the cost associated with delivery time.

Because the optimal delivery day's function is not symmetrical, the loss is different from the Taguchi loss function.

4.4. Likelihood for making a type II error for A

In this case (delivery time for technology), a type II error occurs when, in reality, a product is delivered late but the company thinks the technology item arrived on time.

The likelihood of making a type II error in this case is 0.487613 = 48.7613%

An acceptable type II error ranges around 20%, therefore the type II error is not preferable.

5. Part 5: DOE and MANOVA

For all the hypothesis tests done a p-value of 0.05 was used because it is statistically significant and universally popular to use.

5.1 Hypothesis one

- <u>Dependent variables:</u> Day, Month, Year
- <u>Independent variable:</u> Why bought
- H0: The day, month and year of sales made has no influence on the buying pattern for the reason why products are bought
- H1: At least one feature (Day, Class or Year) has an impact on the buying pattern for why products are bought.

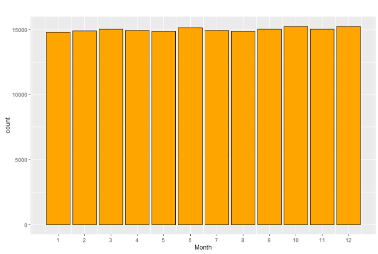
<u>MANOVA test:</u>

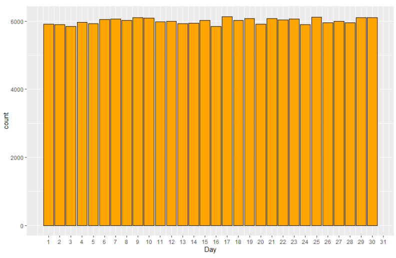
p-value	<2.2e-16	Reject the Null Hypothesis. At least one feature (Day, Class or Year) has an impact on the buying pattern for why products are bought.

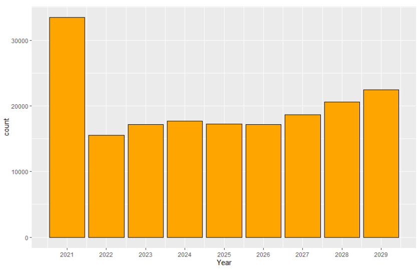
<u>Each dependent variable and class:</u>

Dependent variable	p-value	Analysis
Day	0.5585	The p-value is much higher than 0.05. Day has no impact on the pattern buying for why products are bought.
Month	0.7902	The p-value is much higher than 0.05. Day has no impact on the pattern buying for why products are bought.
Year	2.2e-16	The p-value is smaller than 0.05. Year has an influence on the buying pattern for why products are bought.

<u>Graphs to support results:</u>







(Figure 1, 2 and 3)

<u>Conclusion:</u>

The day and month of a sale made has no influence on the reason why products are bought, however, the year a sale is made has an influence on why a product is bought.

5.2 Hypothesis two

- Dependent variables: Day, Month, Year
- Independent variable: Class
- H0: The day, month and year of sales made has no influence on the buying pattern for products being bought in specific classes.
- H1: At least one feature (Day, Class or Year) has an impact on the buying pattern for products being bought in specific classes.

MANOVA test:

p-value	<2.2e-16	Reject the Null Hypothesis. At least one feature (Day, Class or Year) has an impact on the buying pattern for products being bought in specific classes.

Each dependent variable and class:

Dependent variable	p-value	Analysis
Day	0.1766	The p-value is much higher than 0.05. Day has no impact on the pattern buying for products being bought in specific classes.
Month	0.2859	The p-value is much higher than 0.05. Day has no impact on the pattern buying for products being bought in specific classes.
Year	2.2e-16	The p-value is smaller than 0.05. Year has an influence on the buying pattern for products being bought in specific classes.

Graphs to support results:

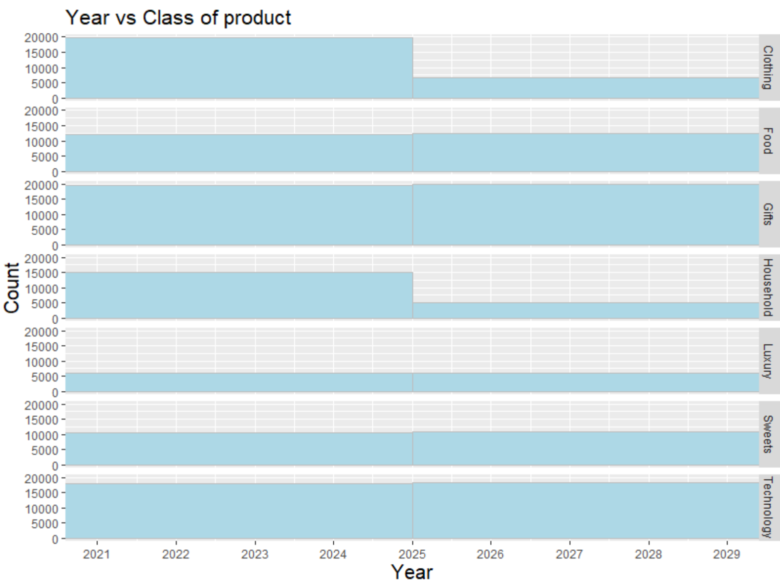


Figure 31: Year vs Class of Product

Conclusion:

The day and month of a sale made has no influence on the class of products sold, however, the year a sale is made has an influence on the class of products sold.

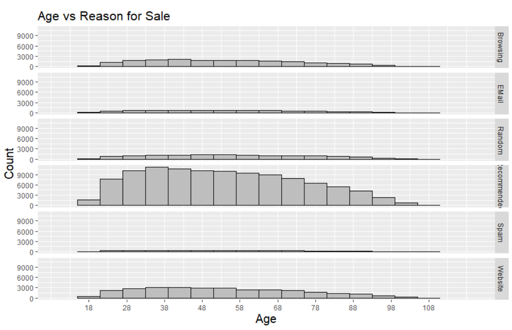
5.3. Hypothesis three

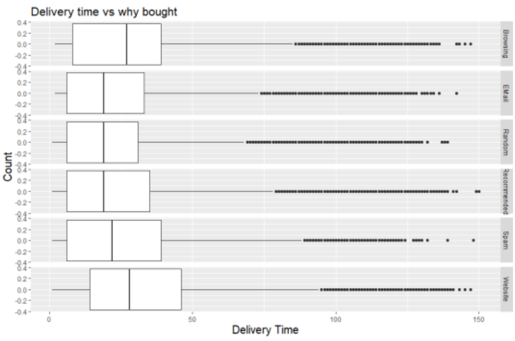
- Dependent variable: Price, Delivery Time, Age
- Independent variable: Why Bought
- H0: The Price, Delivery Time and Age of sales made has no influence on the buying pattern for the reason why products are bought
- H1: At least one feature (Price, Delivery Time, Age) has an impact on the buying pattern for why products are bought.

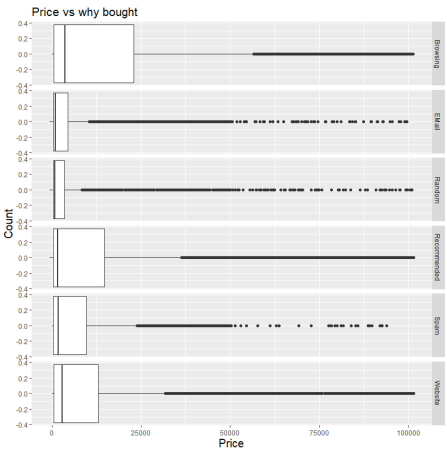
MANOVA test:

p-value	<2.2e-16	Reject the Null Hypothesis. At least one feature (Price, Delivery Time or Age) has an impact on the buying pattern for why products are bought.

Dependent variable	p-value	Analysis
Price	2.2e-16	The p-value is much higher than 0.05. Price has an impact on the pattern buying for why products are bought.
Delivery Time	2.2e-16	The p-value is much higher than 0.05. Delivery Time has an impact on the pattern buying for why products are bought.
Age	2.2e-16	The p-value is smaller than 0.05. Age has an influence on the buying pattern for why products are bought.







Conclusion:

Price, Delivery Time and Age have an influence on the reason why a product is bought.

6. Part 6: Reliability of the service and products

6.1 Problem 6 & 7

Problem 6: A blueprint specification for the thickness of a refrigerator part at Cool Food, Inc. is 0.06 +- 0.04 cm. It costs $45 to scrap a part that is outside specifications. Determine the Taguchi loss function for this situation.

Calculation for k:

$$L(x) = k(x - T)^2$$

$$45 = k(0.04)^2$$

$$k = 28125$$

Taguchi Loss Function:

$$L(x) = k(x - T)^2$$

$$L(x) = 21825(x - 0.06)^2$$

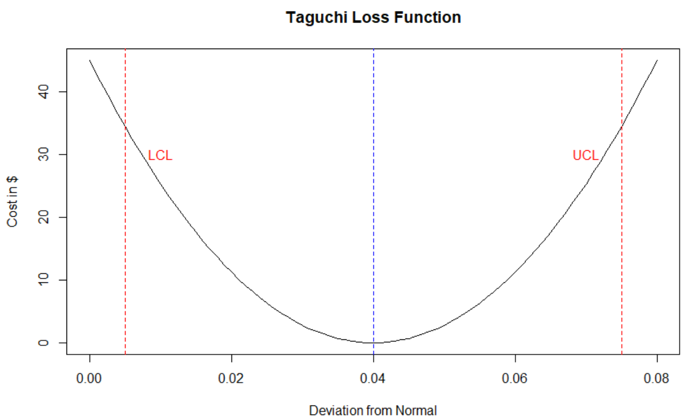


Figure 32: Taguchi Loss Function

When a product's characteristics deviate from the target value (0.04), the quality of the product gets worse. This means that the company will make a loss and that the company will make a loss.

Problem 7: A team was formed to study the refrigerator part at Cool Food, Inc. described in problem 6. While continuing to work to find the root cause of scrap, they find a way to reduce the scrap cost to $35 per part.

a. Determine the Taguchi loss function for this situation
b. If the process deviation from target can be reduced to 0.027 cm, what is the Taguchi loss.

Answer to (a)

Calculation for k:

$L(x) = k(x - T)^2$

$35 \; = k(0.04)^2$

$k = 21875$

Taguchi Loss Function:

$L(x) = k(x - T)^2$

$L(x) = 21875(x - 0.06)^2$

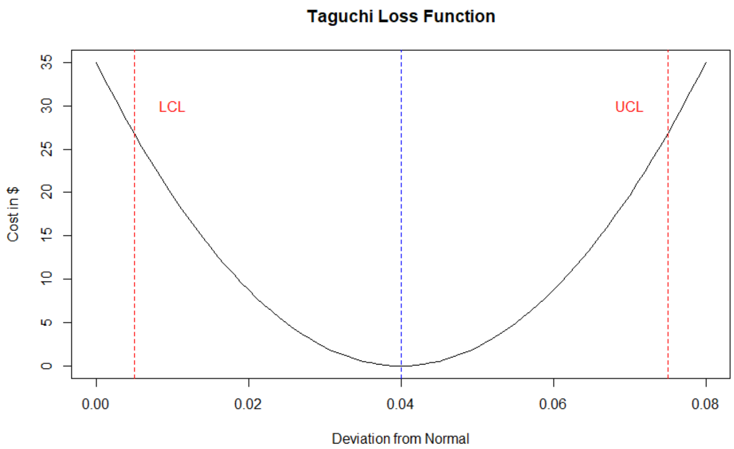


Figure 33: Taguchi Loss Function

When a product's characteristics deviate from the target value (0.04), the quality of the product gets worse. This means that the company will make a loss and that the company will make a loss.

Answer to (b)

The process deviation from target reduced to 0.027cm

$L(x) = k(x - T)^2$

$L(0.027) = 20408.1633(0.027)^2$

$L(0.027) = \$14.87755$

This indicates that the company makes a loss of $14.87755 when the process deviation from target is reduced to 0.027cm.

6.2 Problem 27

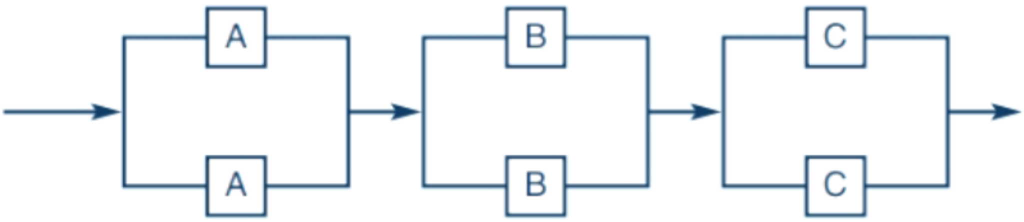


Figure 34: Magnaplex Inc, Production System

Given:

- Reliability of machine A = 0.85
- Reliability of machine B = 0.92
- Reliability of machine C = 0.90

a. <u>System Reliability:</u> of one working machine at stage A, B and C

$$Reliability = Reliability(Machine\ A) \times Reliability(Machine\ B) \times Reliability(Machine\ c)$$

$$Reliability = 0.85 \ \times 0.92 \ \times 0.90$$

$$Reliability = 0.7038 = 70.38\%$$

b. <u>System Reliability:</u> of two parallel machines at stage A, B and C

$$Reliability = Reliability(Machines\ A) \times Reliability(Machines\ B) \times Reliability(Machine\ c)$$

$$Reliability = \big(1 - (0.15 \times 0.15)\big) \times \big(1 - (0.08 \times 0.08)\big) \times \big(1 - (0.1 \times 0.1)\big)$$

$$Reliability = 0.9615 = 96.15\%$$

6.3 Expected reliable delivery times

6.3.1 Part 1

The probability of having reliable vehicles is calculated with the following formula:

$$P(x) = \binom{n}{x} p^x q^{n-x} \frac{n!}{(n-x)!\,x!} p^x q^{n-x}$$

1. $p = 0.03280011$
2. $p = 0.0348579$
3. $p = 0.02701039$
4. $p = 0.02812168$
5. $p = 0.03740828$

Weighted average=

$$\frac{0.0328011(1344) + 0.0348579(190) + 0.02701039(22) + 0.02812168(3) + 0.03740828(1)}{1560}$$

$= 0.03296304$

Expected reliable delivery days in a year (vehicles):

$P(x < 2) = [25C2 \times 0.03296304^2 \times (1 - 0.03296304)^{20-2}] = 0.9731609$

$thus, number\ of\ reliable\ delivery\ days\ in\ a\ year =$

$0.9731609\ \times 365 = 355.2037\ days$

The probability of having reliable drivers is again calculated with the following formula

$$P(x) = \binom{n}{x} p^x q^{n-x} \frac{n!}{(n-x)!\,x!} p^x q^{n-x}$$

1. $p = 0.07408802$
2. $p = 0.08079423$
3. $p = 0.05401372$
4. $p = 0.05498461$

Weighted average=

$$\frac{(1458 \times 0.07408802) + (95 \times 0.08079423) + (6 \times 0.05401372) + (1 \times 0.05498461)}{1560}$$

$= 0.07440696$

Expected reliable delivery days in a year (drivers):

$P(x < 4) = [21C4 \times 0.07440696 \times (1 - 0.07440696)^{21-4}] = 0.9830675$

$thus, number\ of\ reliable\ delivery\ days\ in\ a\ year =$

$0.9830675\ \times 365 = 358.8196\ days$

<u>TOTAL expected reliable days in a year:</u>

Total reliable probability:

$Total\ probability = \ P(vehicles) \times P(drivers)$

$Total\ probability = 0.9731607 \times 0.9830675$

$Total\ probability = \ 0.9566828$

Total expected reliable delivery days in a year=

$0.9566828 \times 365 = 349.1892$ days

6.3.2 Part 2: Increasing the number of vehicles to 21

Using the same equation

$P(x) = \binom{n}{x} p^x q^{n-x} \frac{n!}{(n-x)!\,x!} p^x q^{n-x}$

New delivery reliability:

1. $p = 0.4946574$
2. $p = 0.3540854$
3. $p = 0.1206958$
4. $p = 0.02605607$
5. $p = 0.003996738$
   $Total = 0.9994914$

New total expected reliable delivery days in a year:

$= 0.9994914 \times 365$

$= 364.8144$

7. Conclusion

The dataset given (clients for an online business) was sorted and cleaned. Thereafter, a good understanding of the data was obtained by construction and analyzing several descriptive graphs.

The X&S Charts gave a good indication on the condition of delivery times of the classes. A conclusion was made that household, luxury and gifts products' delivery times are not in control. Further investigation regarding suppliers and reliable delivery times is needed. The Manova tests also led to the same conclusion.

The probability of making a type I and type II error also indicated that the company should re-investigate on-time deliveries of products.

All-inclusive, the importance of applying explorative analysis for a company was learned.

8. References

(2018). Retrieved from BrainKart.com: https://www.brainkart.com/article/Taguchi-s-Quality-Loss-Function_5287/

Complete Dissertation. (2020). Retrieved from https://www.statisticssolutions.com/free-resources/directory-of-statistical-analyses/manova/

Investopedia. (2022, August 1). Retrieved from Descriptive Statistics: Definition, Overview, Types, Example: https://www.investopedia.com/terms/d/descriptive_statistics.asp

M.K., P. W. (2020). *Conjointly*. Retrieved from Desccriptive Statistics: https://conjointly.com/kb/descriptive-statistics/

PDCA Home. (2019). Retrieved from https://pdcahome.com/english/211/quality-loss-function-taguchi-how-to-control-the-lack-of-quality-in-a-product/