QUALITY ASSURANCE - 344

Abstract

This report is an analyses on data for an online compony and recommendations based on the findings of the analyses.

R Heppell

23538589

Contents

Introduction

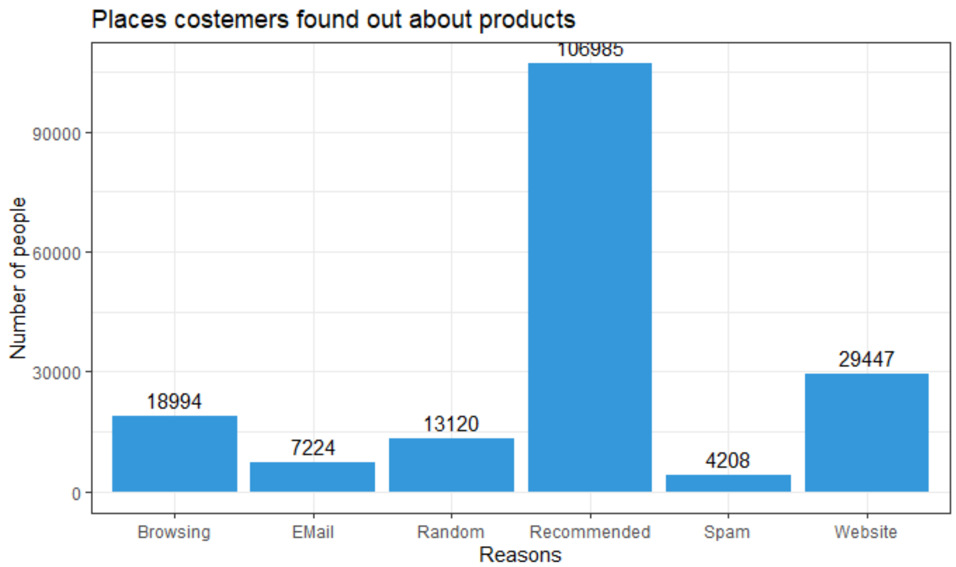
This report will contain the finding made by analysing the data provided of the items that were sold by the online compony. The analyses will focus on the prices of the classes. Further the delivery times will be analysed by making use of statistic process control. Other reliability of aspects will also be investigated. Following the analysis and investigations recommendations will be made. These seek to decrease costs associated with delivery and improve the reliability of the processes.

1 Identifying invalid data

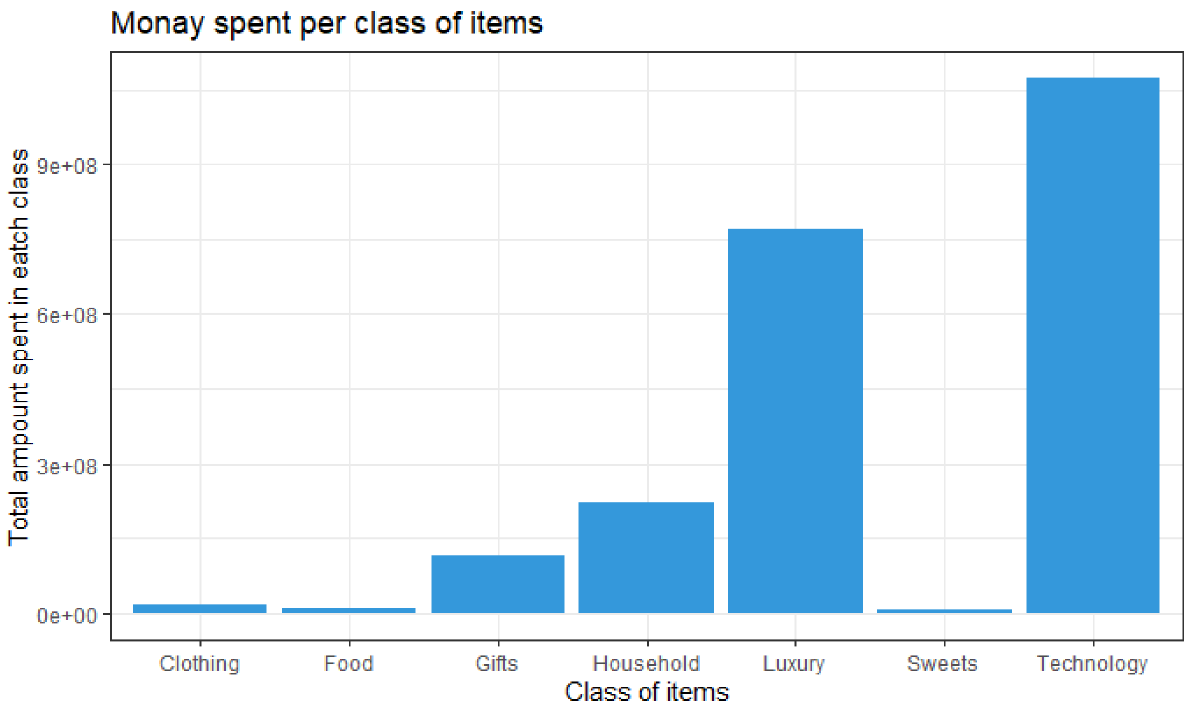
The first step to analysing the sales data was to sort between valid and invalid data. In total there are 180000 instances of data of which 22 instances where classified as invalid. The instances where classified as invalid for not having the price included or for including a negative price. For both of these reasons it could be user error when entering the data or especially for the negative prices it can be an indication of dishonesty. This might be worth looking into at a later stage but the total negative prices entered are less than 0.001% of the total sales.

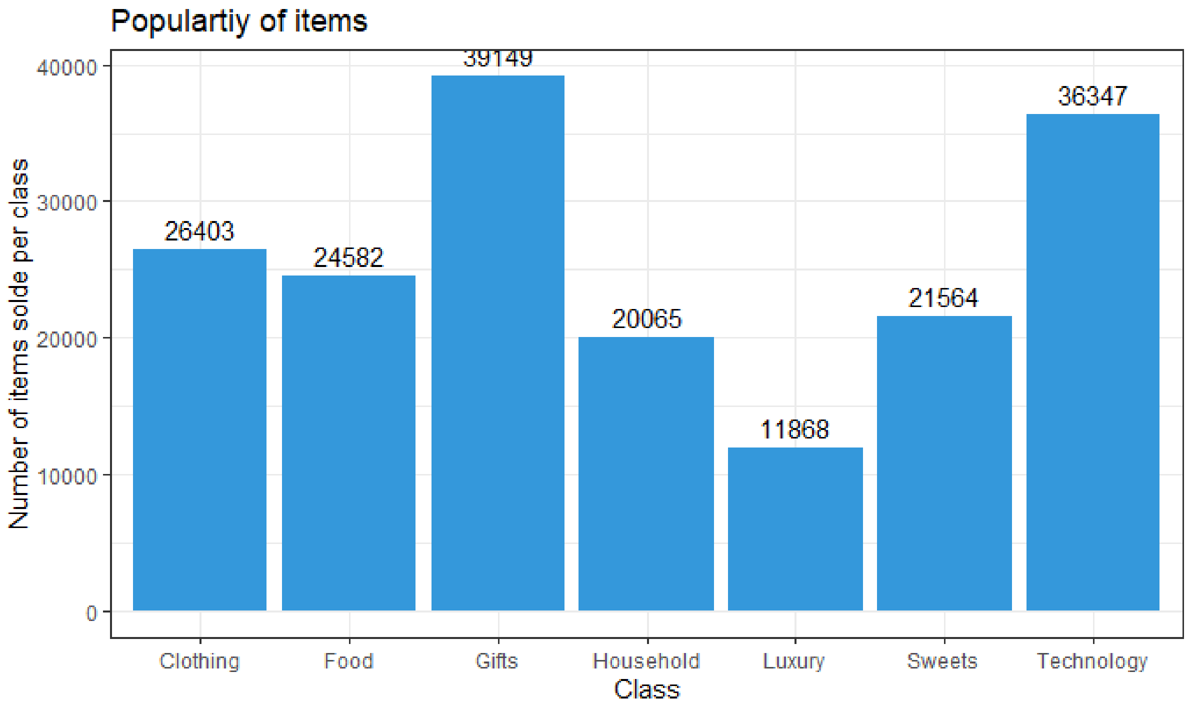
After sorting the data the valid and invalid data was split and saved in 2 separate datasets. From here on only the valid dataset will be considered.

2 Analytical statistics



The above graph indicates the reason that was listed for why a costumer bought an item. The graph clearly indicates that recommendation (word of mouth) was the biggest reason for purchase. From that the assumption can be made that customers are satisfied with the products since they recommended it to others who then became costumers, but this can also indicate that the marketing on the website and email can be improved. The graph also shows that spam generated a verry little customers so might not be worth the resources used for it.





The 2 graphs above show the money spent in each category of items and the amount of items that where sold from each category, respectively. From the data it is clearly visible that the most income is generated by the Technology which is followed by the luxury category. Furthermore the graphs indicate that clothing is the categories that soled the third most items but generated very little income compared to the other categories.

3 Process Capability

For the process capacity the LSL is set to 0 since the perimeter that is used in the calculations is time and it is not possible to have a negative time. In this scenario a negative time would indicate that the item was delivered before the order was placed which is not possible and thus the LSL is chosen to solve the problem.

The process capability indices for the delivery time regarding technology is:

Cp = 504.287

Cpu = 41.90887

Cpl = 210.2346

Cpk = 41.90887

4 Statistical Proses Control

Statistical Proses Control is the use of statistical techniques to manage, control and keep track of a process. (Hessing, n.d.)

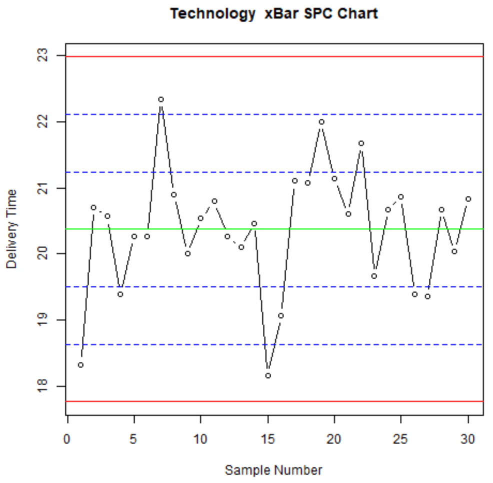
For the analyses of the data for the client`s online business x-bar and s charts where chosen since the delivery time is continues and sample sizes where 15. The analyses was done on the delivery times of the items and divided into sub groups by the class of item. The following values where calculated using the first 30 samples:

4.1.1 xbar values

	LCL	L2Sigma	L1Sigma	CL	U1Sigma	U2Sigma	UCL
Clothing	8.5351	8.68	8.825	8.97	9.115	9.26	9.4049
Household	42.8761	44.1048	45.3335	46.5622	47.7909	49.0196	50.2483
Food	2.2705	2.3437	2.4168	2.49	2.5632	2.6363	2.7095
Technology	17.7743	18.641	19.5077	20.3744	21.2412	22.1079	22.9746
Sweets	2.0585	2.1983	2.338	2.4778	2.6175	2.7573	2.897
Gifts	7.2337	7.6095	7.9853	8.3611	8.7369	9.1127	9.4886
Luxury	3.9771	4.2299	4.4828	4.7356	4.9884	5.2412	5.494

The table above shows the initial values related to the x-cart. These values where calculated on the mean of the first 30 samples (consisting of 15 instincies each). During the processing of the first 30 samples the proces was closely monitord to ensour that no external factors inflowinced the process and coused the dilivery time sot be either shorter or longer.

In the table the Controul Limit (CL) values rever to the target time for delivering a sarten item. The Lower Controul Limit (LCL) is 3 standerd deviations below the CL and the Upper Controul Limit (UCL) 3 standerd diviations ubove the CL. These values along with the mean of every sample in a claas can be used to plot xBar SPC chart.

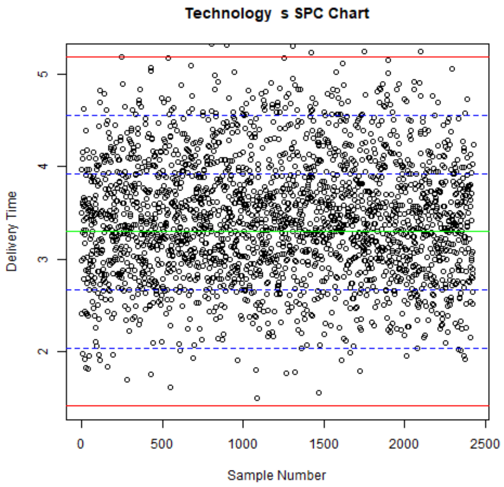


The xBar SPC chart above is the inisilization of the xBar chart for the Technoligy class. The 2 red lines represent the LCL and UCL respectively. The green line represents the CL and the blue dash lines represent the 1 and 2 sigma values. This chart is only the initilisation and thus has only the mean of the first 30 samples ploted. This grape is used in conjunction with the s-chart (below) to identify samples that form paters or represent a proces that is out of controul in other ways.

4.2.2 s values

	LCL	L2Sigma	L1Sigma	CL	U1Sigma	U2Sigma	UCL
Clothing	0.2359	0.341	0.4461	0.5512	0.6564	0.7615	0.8666
Household	1.9996	2.8903	3.7811	4.6719	5.5626	6.4534	7.3442
Food	0.119	0.1721	0.2251	0.2781	0.3312	0.3842	0.4372
Technology	1.4105	2.0388	2.6672	3.2955	3.9239	4.5522	5.1806
Sweets	0.2274	0.3288	0.4301	0.5314	0.6327	0.734	0.8353
Gifts	0.6116	0.8841	1.1565	1.429	1.7014	1.9739	2.2463
Luxury	0.4114	0.5947	0.778	0.9612	1.1445	1.3278	1.5111

This table represent the s SPC Chart values that was calculated on the same first 30 sample as the xBar SPC chart values. For the s chart the values are calculated on the standard deviation of each class rather than the mean as is used for the xBar chart. Once again the CL represent the trget to strive for in each class.



The s SPC Chart above is a representation of the standard deviation of each sample in the technology class. The red lines represent the UCL and LCL and the green the CL as in the xBar chart. The values represented by the different lines are calculated on the first 30 samples an then only changes if the process itself changes. For instance a new delivery vehicle is added to the fleet. Then the standard deviation for each sample is plotted and as is visible on the graph there are a view samples where the standard deviation of that sample is not within the control limits. This is an indication that the process is out of control and would weren't an investigation to find out why and correct the fault causing the discrepancy.

4.2 Type 1 Errors

A type 1 error (Manufacture's) error occurs when the process is in control but the signs point to the process being out of control. This scenario would result in an investigation and possibly a change made to the process when the process was truly in control.

The likely hood of making a type 1 error (Manufacturer's) error is 0.0027 or 0.27%. This is because the data is normally distributed and 99.73% of the data is whit in 3 standard deviations of the mean in either direction. This is visually represented by the LCL and UCL on the previous charts.

From that it is assumed that reacting to a signal indicating the process is out of control (a sample mean being above or below the UCL or LCL respectively) there is a 0.0027 possibility of making a mistake. More importantly not reacting to a signal showing the process is out of control has a likely hood of 0.9973 of being a mistake.

4.2 Our of Control xBar

	1	2	3	4	5	6	Number out of s
Clothing	8.5333	8.5333	8.5333	9.4333	9.4667	9.5	17
usehold	42.2333	42.4667	42.8333	57.3667	54.5667	55.8	400
Food	2.2667	2.2667	2.2667	2.2667	2.7333	NA	5
hnology	17.5	17.2667	17.4333	16.9667	17.5	23.2667	17
Sweets	2.0333	2	2.9	2.9667	2.9667	NA	5
Gifts	10.2333	9.6667	9.7	16.5667	16.3	16.0333	2290
Luxury	3.9667	3.9667	3.9	3.4	3.3	3.6	434

This table represents the xBar values that were out of control for each class. In this case the instances where only deemed as out of control if they were above the UCL or below the LCL and not taking in to account all the other rules. In the table the first three columns represent the first three values that were out of control in each class. The next three columns represent the last three instances that were out of control except for the Food and Sweets classes which only had 5 instances that were out of control and all 5 are shown. The last column shows the total number of instances that were out of out of control in each class.

	Longest run in controul	Last ample
Clothing	4	0.5345
Household	3	4.5638
Food	7	0.2673
Technology	6	3.2671
Sweets	4	0.5071
Gifts	5	1.4125
Luxury	4	0.9537

This table represents the larges continues number of instances that were les than 0.3 sigma below the CL an less than 0.4 sigma above the CL and thus being very close to the target. The second colom represents the value od the last instance of the longest number of instances in between the bounds for that particular class.

4.3 Cost saving on delivery

As the satiation currently is, 1356 items from the technology class was delivered late (having a delivery time of more than 26 hours). These late deliveries resulted in a "cost" of R12 357 898.00. By increasing the mean delivery time by 6 hours the total "cost" due to late deliveries of items in the technology the "cost" can be reduced to R 599 161.00. By implementing the change to increase the mean delivery time by 6 hours at R2.5 per item per hour the total savings will be R11 758 737.00. The number of items being delivered late will be reduced to only 5 items.

5Taguchi loss function

The Taguchi loss function states that a product is not suddenly worthless as soon as it is outside of the acceptable bounds but rather that the customer loses value as te product moves away from the initial design specifications.

When a part cost $45 to scrap and the dimensions are 0.06+- 0.04 the Taguchi loss function is:

$L = 28125(y-0.06)^2$

When the cost of scrapping a part is reduced to $35 the Taguchi loss function is:

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

This function represent the value in $ that this compony (not Lafrideradora) will lose in relation to the true dimension of the part. This true dimension is represented by y in cm.

Thus if the cost to scrap remains $35 the loss incurred from a part of 0.027cm is:

$L= 21875*0.027^2 = 15.95$

This means that for a part with a thickness of 0.027cm it will not be scrapped but the this company will incur a loss in value equal to $15.95.

6 System reliability

The current process has identical machines to have one working and one machine as a backup. This can seem wasteful as the backup machines means that the capacity of the process is double of what it produces. This can be bad as it is capital that can not be used elsewhere and is generating little to no value.

Firstly this can be justified by being able to perform regular scheduled maintenance on the machines without interrupting the process. The backups also decreases the possibility of having down time and not being able to complete the process.

For instance if there was no backup machines the reliability of the system would be:

0.85*0.92*0.9 = 0.7038

This means that there is only a 70% chance of the system being able to run fully. That leaves a roughly 30% chance that at any given moment the system will not be able to run and production would instantly halt. Such a holt in production would result in major losses.

Were as with the backup machines the likely hood of the process running smoothly is:

0.9775*0.9936*0.99 = 0.9615

This means that there is less than 4% chance that the process will not run.

By having the backup machines the reliability increases by roughly 26% and considering the losses that would be incurred due to the process being halted the extra capital is worthwhile being tied up in the backup machines.

Conclusion

On completion of the analyses there where now inherent flaws idented but areas for improvement where identified. If the number of instances with invalid prices increases to above 2% it would be worth investigating for fraud or flaws in the current system. The Luxury and technology classes provide the largest income and should thus be receive more recourses than other classes providing less income. The samples in the technology class that where identified as out of control should be investigated and rectified. The loss from on parts for refrigerated delivery vehicles is within reason and does not pose a problem as of yet. The redundancy provided by the backup machines are important and not a waist of recourses and is working well as is.

Bibliography

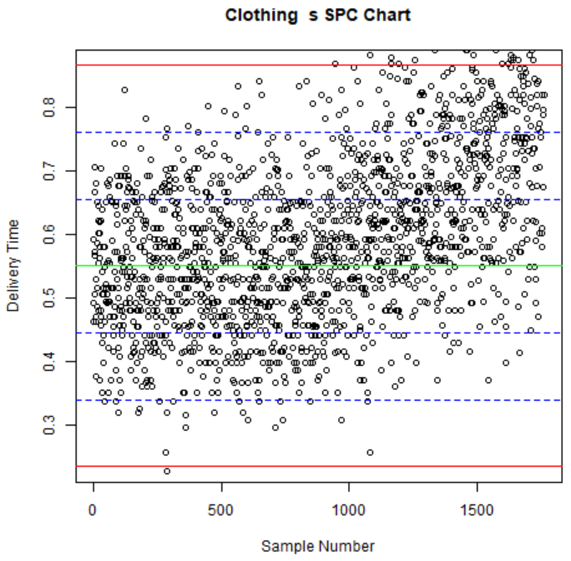
Hessing, T. (n.d.). *sixsigmastudyguide.com*. Retrieved from Statistical Process Control (SPC): https://sixsigmastudyguide.com/statistical-process-control-spc/

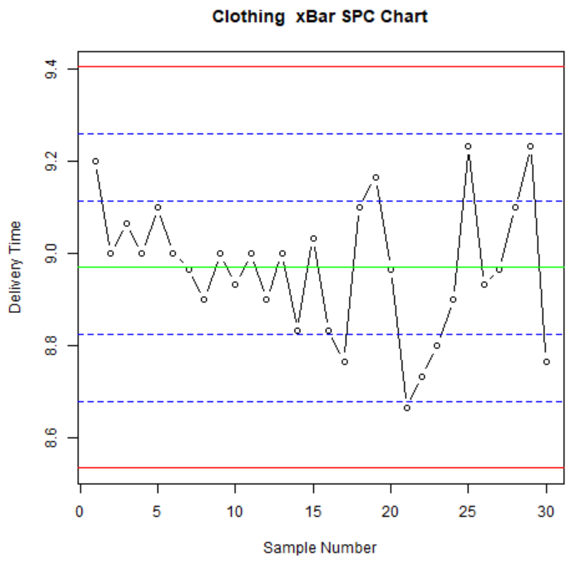
Oakland, J. S. (2007). Statistical Process Control. In J. S. Oakland, *Statistical Process Control* (p. 472). London: Routledge.

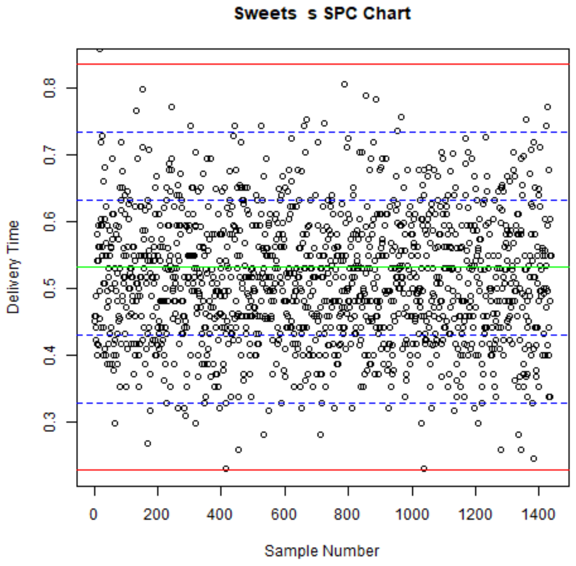
Wachs, S. (n.d.). *Where do the typical control chart signals come from?* Retrieved from winspc.com: https://www.winspc.com/where-do-the-typical-control-chart-signals-come-from/#:~:text=By%20convention%2C%20the%20probability%20of,being%20plotted%20is%20normally%20distributed.

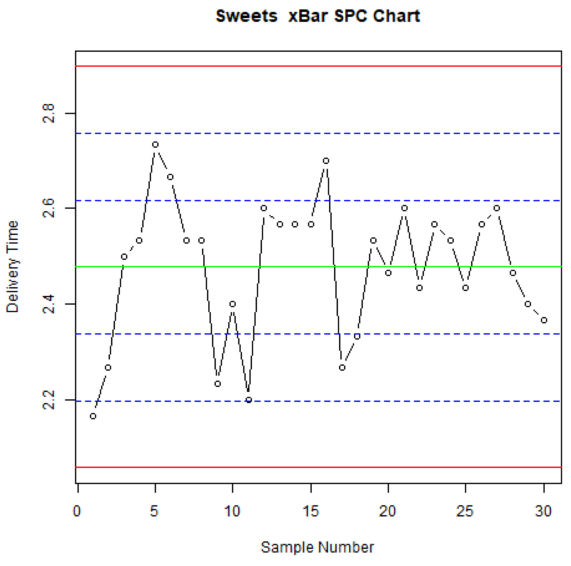
Appendix A

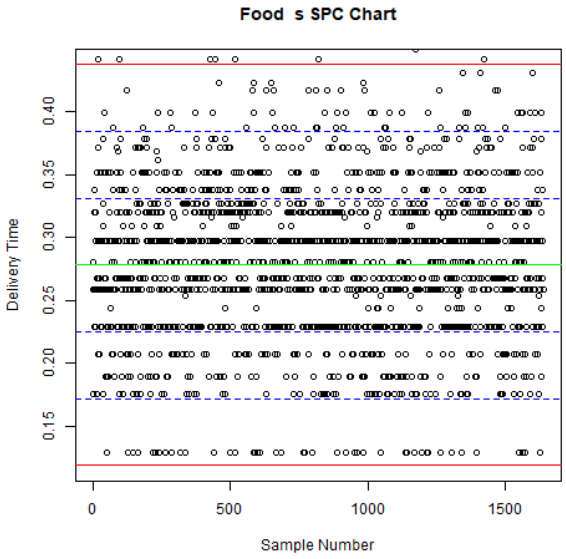
xBar an s SPC charts for all categories

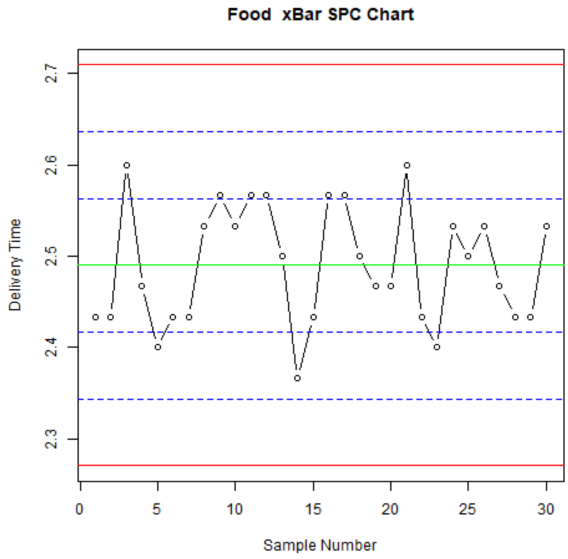


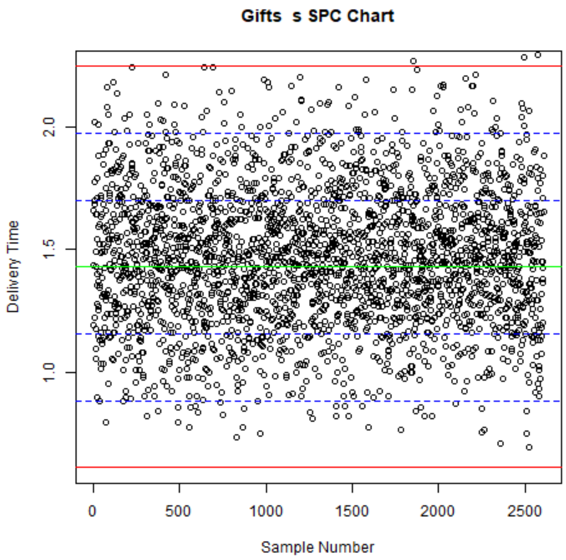


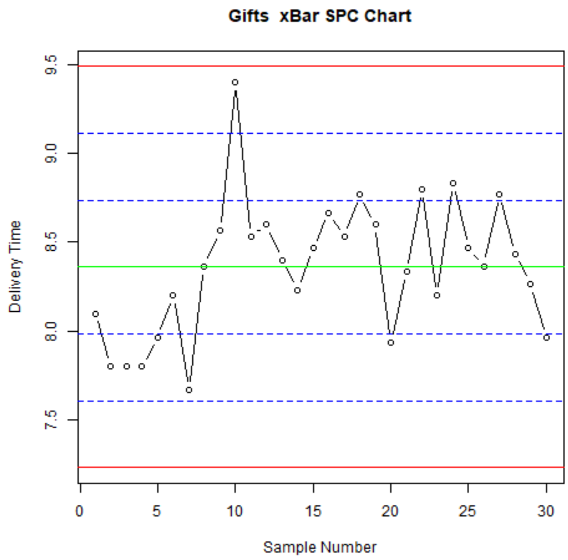


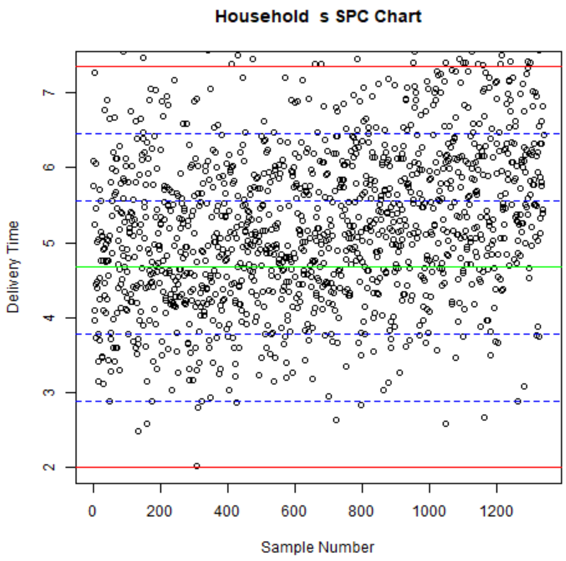


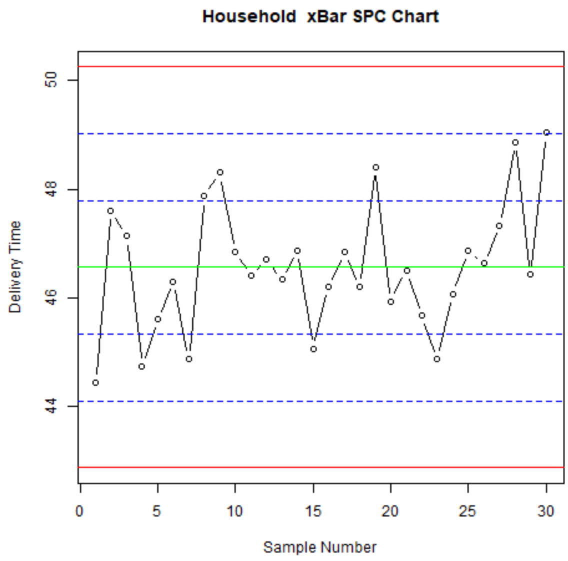

Food s SPC Chart
Delivery Time
Sample Number

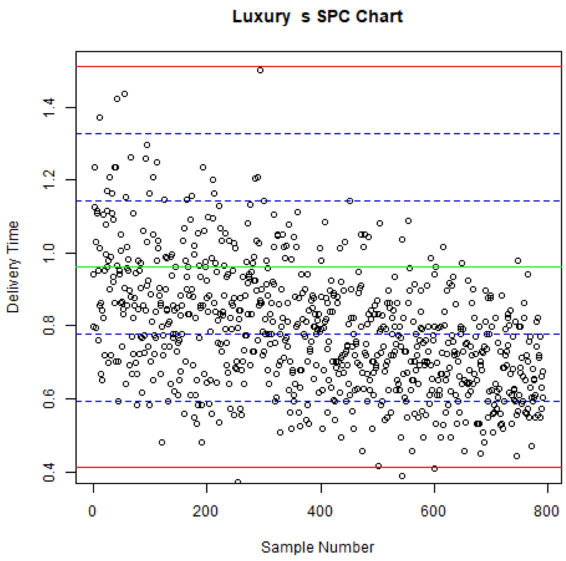

Food xBar SPC Chart
Delivery Time
Sample Number

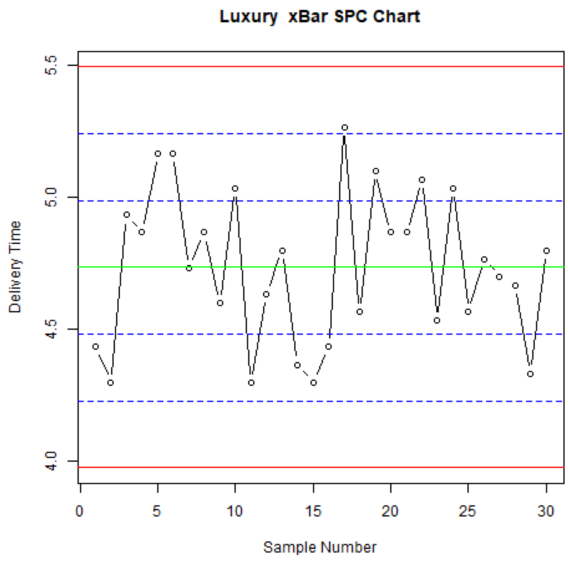

Gifts s SPC Chart
Delivery Time
Sample Number


Gifts xBar SPC Chart
Delivery Time
Sample Number


Household s SPC Chart
Delivery Time
Sample Number


Household xBar SPC Chart
Delivery Time
Sample Number


Luxury s SPC Chart
Delivery Time
Sample Number


Luxury xBar SPC Chart
Delivery Time
Sample Number

V