Paper 088-2013

Demand Forecasting Using a Growth Model and Negative Binomial Regression Framework

Cally Ong Yeru, Singapore Management University
Murphy Choy, Singapore Management University
Michelle Cheong, Singapore Management University

ABSTRACT

In this paper, we look at demand forecasting by using a growth model and negative binomial regression framework. Using cumulative sales, we model the sales data for different wristwatch brands and relate it to their sales and growth characteristics. We apply clustering to determine the distinctive characteristics of each individual cluster. Four different growth models are applied to the clusters to find the most suitable growth model to be used. After determining the appropriate growth model to be applied, we then forecast the sales by applying the model to new products being launched in the market and continue to monitor the model further.

INTRODUCTION

Fashion is a general term for popular style or practice, especially in clothing, footwear, accessories, makeup or furniture. Fashion usually refers to the newest creations made by designers and are bought by consumers. At the same time, fashion is constantly changing with time. One trend can be popular today and out the next hence it is important that right products are being introduced into the market at the right time.

The fashion industry consists of four levels: the production of raw materials; the production of fashion goods by designers, manufacturers and/or contractors; retail sales and last but not least various forms of advertising and promotions. These levels consist of many separate yet interdependent sectors, all of which are dedicated to the goal of satisfying consumer demand under conditions that enable participants in the industry to operate at a profit.

Growth curve analysis is perhaps the most common kind of analysis to be associated in the retail industry. In term of techniques applied, clustering is the most commonly applied exploratory method in the data mining process, where it is used for data exploration and in conjunction with other data mining techniques such as regression or classification. The two most commonly used clustering techniques are Partitional clustering and Hierarchical clustering. Partitional clustering tend to minimize cost function or an optimization criterion, associating a cost to each assignment of an object to a group whereas Hierarchical clustering enable better versatility as they do not require a clear definition of the number of clusters to find.

In this paper, we start by exploring the retail industry outlook to understand the consumer purchasing behaviour. By better understanding the consumer purchasing behaviour, retailers will be more aware of the consumers' needs hence more effective strategies can be devised. As the analysis is more focused on the watch industry, we will also look into details the watch industry outlook.

Having mentioned that Growth curve analysis is the most common used analysis, we will implement cumulative sales data of a retail company to the Growth curve analysis to understand the sales pattern of their products to give the retailer on when new products should be introduced into the market.

RETAIL INDUSTRY OUTLOOK

With the reduction in consumer spending and poor consumer credit market, the retail industry outlook remains as challenging. With more informed buyers who are more aware of their needs and desire, retailers need to deliver an enthralling in-store experience in order to drive profitable growth in the coming periods.

In order to stay competitive, retailers would need to devise strategies that align talent, physical space, processes and technology to meet the changing demands of consumers. This can be achieved by having an integrated inventory system that provide transparency to product availability, targeted training programs for the retail associates and implemented digital technology that deliver a more immersive customer experience.

Analytics has always been one of the most important and powerful tools available for retailers. It helps retailers to better understanding on customers' needs and behaviour and helps retailers to create targeted promotion to entice customers to purchase which ensure marketing investment are done effectively. In addition, it benefits retailers by allowing them to balance their inventory with the customer demands with the use of analytics.

Having a clear visibility of the products, retailers will be able to offer desirable goods and services that consumer would want to buy. By effectively managing inventory, retailers will be able to explore new markets to extend the reach of their products.

WATCH INDUSTRY OUTLOOK

With the increase in luxury watch purchases, a growth of +12% vs. 2009 of retail and jewellery sales was observed in the first half for 2010 and remain stable towards the end of 2010. As of December 2011, Singapore had around 304 stores selling watches and out of which less than 15% was selling luxury brand watches. Currently, Singapore was ranked 7th worldwide in terms of Swiss watch imports which grew by more than 33% from 2010 same period. Singapore was the leading market in Southeast Asia which is followed closely by Thailand and Malaysia.

A survey done in December 2011 on watch retailers showed that more than 30% of their customers were tourists with a majority coming from China and India. The reason of this high percentage was largely contributed due to their willingness of purchasing luxury watches owing to its authenticity as it is less likely of them getting fakes. The other contributing factor was due to the availability of tax refund hence paying lesser than in their countries where such products are heavily taxed. However, in 2012 a slight decline (-2% vs. 2011) in watch and jewellery sales was observed.

With the decline in watch and jewellery sales, this pushes retailers further to devise effective strategies to maintain a steady growth in sales when new products are being introduced into the market. At the same time being a data-intensive industry, retailers will need to be able to make use of the data to better operate and manage their business with analytics which we will discuss in the following sections.

GROWTH MODEL AND NEGATIVE BINOMIAL REGRESSION

GROWTH MODELS

Growth model is a technique that statistically assesses change and correlates of change. With the understanding that time series is a sequence of observations that are arranged according to the time of their outcome, four growth models were applied to understand which growth curves will be most applicable.

Logistic Function

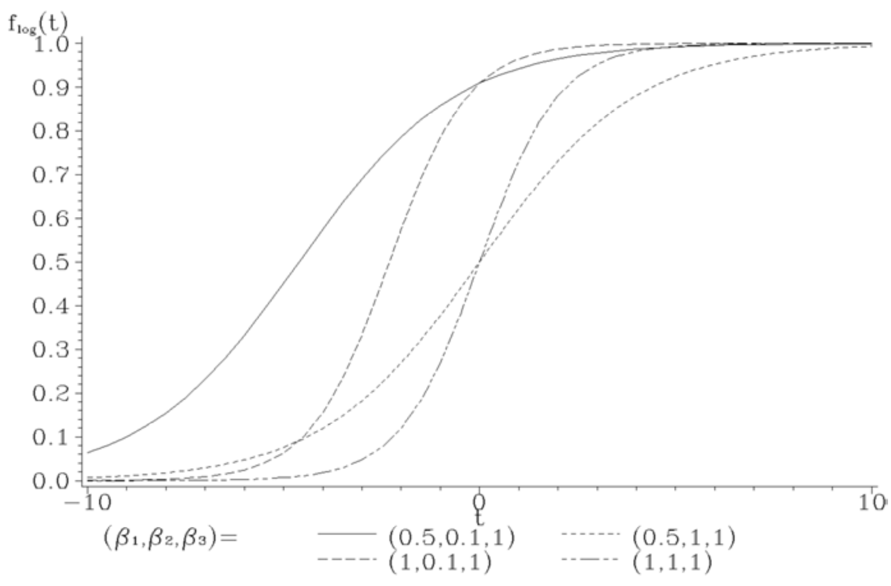


Diagram 1: Logistic Function (Falk et. al., 2006)

The function $f_{log} := \frac{\beta_3}{1+\beta_2 \exp(-\beta_1 t)}$ is a widely used logistic function however there is a limit for $f_{log}(t) = \beta_3, if\ \beta_1 > 0$ where β_3 is similar to the maximum impregnation or growth of the system. In addition, a linear relationship is represented among $\frac{1}{f_{log}(t)}$; hence we can use this as the basis for estimating the parameters of $\beta_1, \beta_2, \beta_3$ with an appropriate linear least squares approach.

Mitscherlich Function

Mitscherlich function is another function that is usually used to model a long term growth of a system. As the function indicates:

$$fM\ (t) \coloneqq fM(t, \beta_1, \beta_2, \beta_3) \coloneqq \beta_1 + \beta_2 \exp(\beta_3 t), t \geq 0,$$

where $\beta_1, \beta_2 \in \mathbb{R}$ and $\beta_3 < 0$, there will be the asymptotic behaviour since β_3 is negative and parameter β_1 is the saturation value of the system. The value of the system at the time t = 0 will then be$f_M(0) = \beta_1 + \beta_2$.

The Gompertz Curve

Another common function used to model the increase or decrease of a system is the Gompertz Curve.

$$f_G(t) \coloneqq f_G(t;\ \beta_1, \beta_2, \beta_3) \coloneqq \exp(\beta_1 + \beta_2 \beta_3^t), \qquad t \geq 0,$$

Where $\beta_1, \beta_2 \in \mathbb{R}$ and $\beta_3 \in (0,1)$.

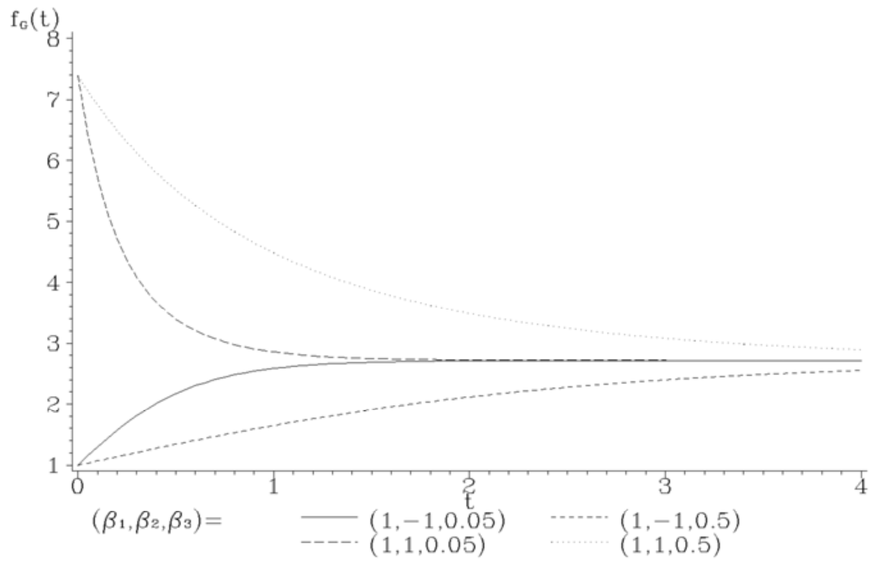


Diagram 2: Gompertz Curve (Falk et. al., 2006)

With the understanding of$\log(f_G(t)) = \beta_1 + \beta_2 \beta_3^t = \beta_1 + \beta_2 \exp(\log(\beta_3)\, t)$, $\log(f_G)$ is a Mitscherlich function with parameters β_1, β_2 and$\log(\beta_3)$. The saturation size obviously is $\exp(\beta_1)$.

The Allometric Function

A common trend function used in biometry and economics is Allometric function,

$f_a(t) \coloneqq f_a(t;\ \beta_1, \beta_2) = \beta_1 t^{\beta_1}, t \geq 0$, with $\beta_1 \in \mathbb{R}, \beta_2 > 0$. This function can also be viewed as particular Cobb-Douglas function where it is a popular econometric model to describe the output produced by a system depending on an input. As

$$\log(f_a(t)) = (\log(\beta_2) + \beta_1 \log(t), t > 0,$$

is a linear function of log(t), with slope β_1 and intercept $\log(\beta_2)$, we can adopt a linear relationship model for the logarithmic data log(y_t)

$$\log(y_t) = \log(\beta_2) + \beta_1 \log(t) + \varepsilon_t, t \geq 1,$$

where ε_t are the error variables.

POISSON DISTRIBUTION

Poisson distribution is used to model information on counts of various kinds, particularly in situations where there is no natural 'denominator', hence there is no upper bound or limit on how large an observed count can be. Some examples of count data where poisson distribution is used are 1) the number of car accidents in a given area over year intervals, 2) the number of H1N1 cases for a given risk group for a series of monthly intervals, 3) the number of robberies in Singapore by year. As mentioned in each of the examples, there is no reasonable denominator associated with the counts.

When a Poisson model is applicable for an outcome Y, the probabilities of observing any specific count, y, are given by the formula:

$$\Pr(Y = y) = \frac{\lambda^y e^{-y}}{y!}$$

where λ is known as the rate parameter, and y! = y x (y-1) x …x 2 x 1. Such a Poisson random variable Y has expectation E(Y) = λ, and variance Var(Y) = λ. The fact that the expectation and variance agree provides a quick check on whether a Poisson model might be appropriate for a sample of observations.

NEGATIVE BINOMIAL REGRESSION

An extension from the Poisson distribution would be Negative Binomial Regression. The maximum likelihood procedure used to derive estimates and provide estimates variability of these estimates in Poisson regression makes a strong assumption that every subject within a covariate group has the same underlying rate of the outcome. This also implies that the variability of counts within covariate is equal to the mean or:

$$var\left(Y\left(X_1 X_2, \ldots, X_p\right)\right) = \exp(a + b_1 X_1 + b_2 X_2 + \cdots + b_p X_p)$$

If this fails to be true, the estimates of the coefficients can still be consistent using Poisson regression, but the standard errors can be biased and they will be too small. We would not expect that we have measured every variable that contributes to the rates of events, so there will always be residual variations in the rates of events amount people who all have the same covariate values. Using the Negative Binomial Regression, it can account for greater than Poisson variation and is based on the negative binomial distribution.

A negative binomial probability of Y is:

$$P(Y = y) = (\tfrac{r}{r+\lambda})^r \, \tfrac{\Gamma(r+y)}{\Gamma(y+1)\Gamma(r)} (\tfrac{\lambda}{r+\lambda})^y,$$

where Γ is the gamma function. The mean of the negative binomial distribution (like the Poisson) is λ but the variance is $\lambda + \lambda^2/r$, where r is called the dispersion parameter. With reference to the **Diagram 3** below, as r gets large (and λ is fixed), the negative binomial converges to a Poisson distribution. This means the negative binomial model is a more general model as compared to the Poisson and it can lead to being as a mixture of Poisson distributions. In addition, testing can be done between negative binomial distribution and Poisson regression by fitting both models and perform a likelihood ratio test to identify which has a more significant improvement over the other.

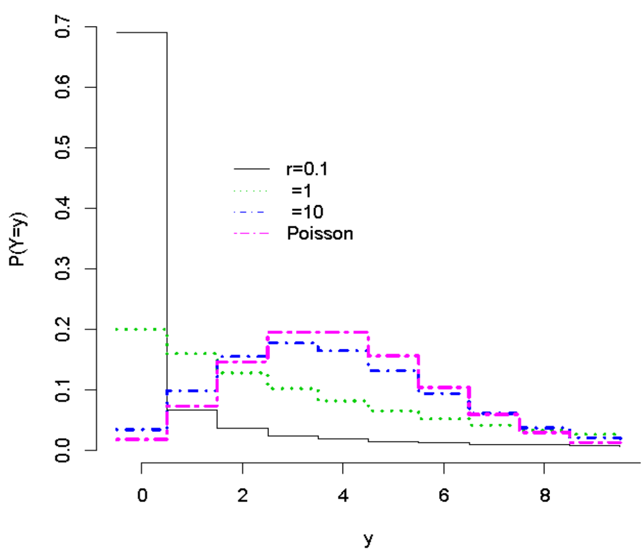


Diagram 3: Negative Binomial Probability Distribution (NICHOLAS P. JEWELL, 2006)

CLASSICAL CLUSTERING

There are 2 types of clustering that are used most often which are partitional and hierarchical clustering.

Partitional Clustering

Partitional clustering algorithms tend to minimize a cost function or an optimization criterion, associating a cost to each assignment of an object to a group. By often, the partitional clustering usually forces the algorithm to search for the most optimal solution by starting from a random initial solution and evolving at each iteration, until the optimality condition is met.

The K-means is classified as one of the methods used in partitional clustering where each group is represented by the gravity centre of the group. It is probably one of the first clustering procedures which converge into a solution starting from an initial set of *K* cluster centres, assigning each point to the closest group in Euclidean distance after which the centroids of the groups are replaced with the mean points of the corresponding group.

K-means is a heuristic solution to the clustering problem based on the assumption that the examples are produced by *k* spherical Gaussian distributions. The aim is to produce *k* groups from *n* given objects in order to minimize the quadratic error. Having *p* be each of the objects belonging to group C_i, m_i the mean of objects within that group. Assuming that all *n* objects are assigned to a specific group, the quadratic error is given by

$$EQ = \sum_{i=1}^{k} \sum_{p \in C_i} |p - m_i|^2$$

The mean of a specific group is a vector with the mean values of each attribute for all objects included in that group. This requires *k* (number of groups) as an input parameter and produces as a result the mean points of each group C_i, with i = 1,…, *k*. Hence, the algorithm is described as per below:

1. Choose *k* objects as initial means
2. Assign each one of the *n* objects to the closest centroid
3. Recompute the means of the *k* groups
4. Repeat steps 2 and 3 while the centroids changed

Hierarchical Clustering

Hierarchical algorithms set the basis on dissimilarities among elements of the same group by defining a hierarchical structure of clusters. Agglomerative hierarchical methods start with as many clusters as initial objects and build the structure by clustering the two most similar groups at each step. Divisive hierarchical methods start with one large cluster containing all objects, and build the structure by dividing the group with the largest diameter in two new groups.

There are three advantages of using hierarchical clustering as compared to partitional clustering where they do not require user-defined number of target clusters and it does not make any assumptions about data distribution. At the same time, no explicit representation apart from pair-wise dissimilarity matrix is required for hierarchical clustering. It is also noted that non parametric algorithms set their basis on dissimilarities between elements of the same cluster or between clusters, defining a hierarchy of clusters, either in an agglomerative or divisive way.

A common feature of all algorithms in this area is the use of a dissimilarity matrix. This matrix can be implemented as a vector diss with size $\frac{n(n-1)}{2}$, where the dissimilarity between objects x_i and x_j, with $1 \le i < j \le n$, is saved at

$$d(x_i, x_j) = diss\,[n * (i-1) - i * \frac{(i-1)}{2} + j - i]$$

There are no restrictions on the measure used to compute the dissimilarities and it should comply with the distance metrics' features. However, when the dataset size gets large, calculation of the dissimilarity matrix can prove to be expensive. Hence, for hierarchical models, split at each level of the hierarchy can be based on the k-means techniques or Expectation Maximization.

One of the hierarchical clustering methods used is DIANA, **DI**visive **ANA**lysis where it is a divisive hierarchical clustering system that builds a hierarchy of groups using a top-down strategy. At each step, the cluster with higher diameter (the diameter measure used is the highest dissimilarity between objects within the cluster) is divided into two, until all groups contain only one object, executing n-1 iterations. There are $2^{n-1} - 1$ possibilities of dividing data in two groups on divisive clustering.

At the end of the process, it will always split the data into singular groups and yet includes the divisive coefficient, a measure of quality on the performed divisions. The strength of clustering structure is represented by this coefficient as found by the algorithm. It is also typically used for comparing two hierarchical structures achieved with different algorithms or parameters. The divisive coefficient of a given clustering structure for n objects is

$$DC_{clust} = 1 - \sum_{i=1}^{n} \frac{dd(i)}{n}$$

where *dd(i)* is the quotient between cluster's diameter to which object v_i belonged before being split into singular group and the global diameter of the dataset. DIANA clustering algorithm is described as below:

1. Select the group C_k with the largest diameter
2. Let $D^k = \{d(x_i, x_j)|x_i, x_j \in C_k\}$ and the subset for a fixed x_i given by $D_i^k = \{d(x_i, x_j)|x_i, x_j \in C_k\{x_i\}\}$
3. Start a new cluster $C_s = \{x\}$ where $x_s = \underset{x_i}{argmax}(\overline{D_i^k})$
4. Let $D_i^{ks} = \{d(x_i, x_j)|x_i, x_j \in C_k, x_j \in C_s\}$ and $D_i^{kk} = \{d(x_i, x_j)|x_i, x_j \in C_k, x_j \notin C_s\}$
5. $x_h = \underset{x_i \in C_k \setminus C_S}{argmax}(\overline{D_i^{kk}} - \overline{D_i^{ks}})$
6. If $\overline{D_i^{kk}} - \overline{D_i^{ks}} > 0$ then x_h is, on average, closer to the splinter group C_s than to the original one C_k; so move x_h into C_s
7. Repeat steps 4 through 6 until all $\overline{D_i^{kk}} - \overline{D_i^{ks}} < 0$
8. If exist one group with more than one object, go back to step 1

DATA EXPLORATION & WATCH DEFINITION

Five types of watches are generally used to identify the different types of watches in the industry. They are namely Fad, Fashion, Trend, Lifestyle and Brand Presence watches.

Fad represents a temporary fashion that is taken up with great enthusiasm for a brief period time. Fad watches general have a short lifespan of 1 to 3 months where majority of the sales are generated. We can see examples of Fad fashion such as the Korean girl group, Wondergirls or even the infamous planking that happen around the world for a short period of time.

From the analysis, we identify that most of the brands falls under this category. They tend to have a sharp sales increase and flattens quickly once its' hit the peak. An example of a Fad sales growth can be seen in the diagram below.

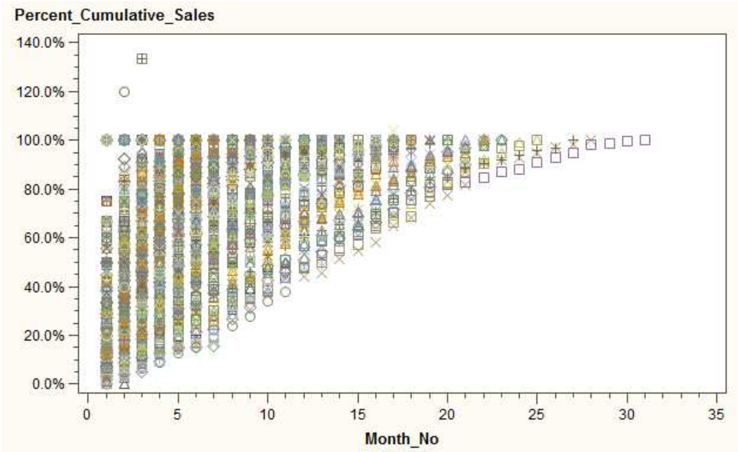


Diagram 4: Fad Watch

Another type of watches that were identified was Fashion watches. Fashion refers to a popular trend where it is constantly changing and has the power to make a social statement. They usually have a lifespan between 3 to 6 months. Examples of fashion that we can relate to are bell bottoms or shiny coloured shoelaces that were popular in the early days.

Trend watches tend to refer to something that is developing or changing in a general direction with a medium lifespan of 6 months. Retro fashion is one of the examples where it was popular for a period of time and several teenagers were dressing up in the 70s and 80s. We can see that Retro theme was a popular theme used in corporate's Dinner and Dance events as well. Trend watches have a more constant sales growth as compared to Fad or Fashion sales growth but will flatten as well towards the end of the cycle. One of such example can be found in the diagram below.

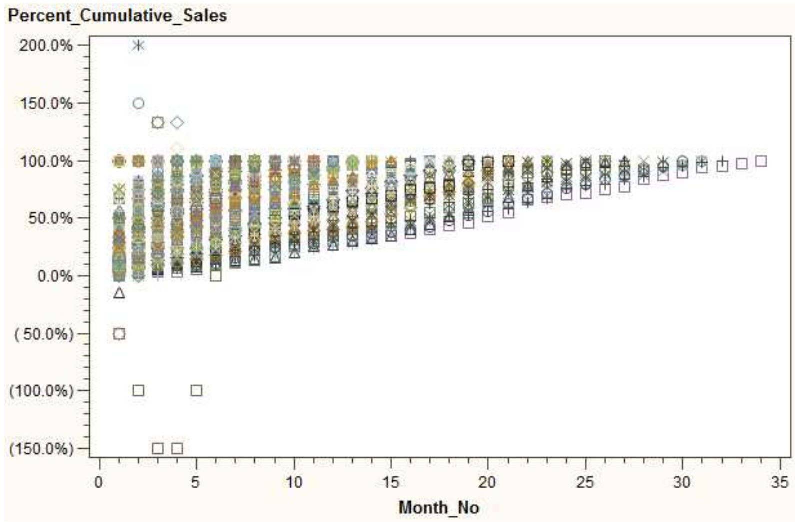


Diagram 5: Trend Watch

A fourth type of watch that was identified from the analysis was Lifestyle watches. Lifestyle is a style that reflects the attitude of a person or group. Such styles usually span around 12 to 18 months which generally last longer as compared to the three types of watches identified earlier. Fashion such as senior executive look can be related as lifestyle fashion where it reflects a more mature composure. Lifestyle watches have a slow sales growth observed during the beginning of the cycle and slowly moving towards an increase in sales as it move towards the end of the cycle as shown below.

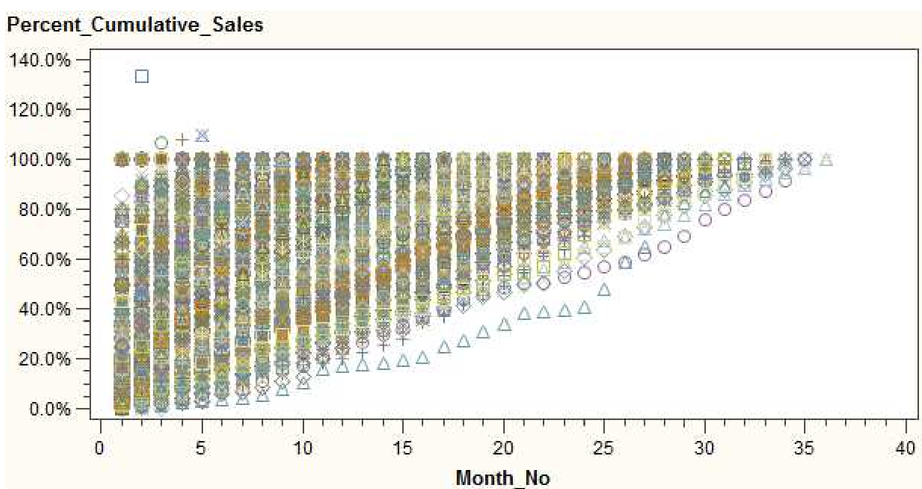


Diagram 6: Lifestyle Watch

Finally, Brand presence watches represent the last type of watch. Style that belongs to the brand presence category is recognized by consumer easily. Luxury brands generally fall in this category and it has a longer lifespan of 18 months or more. Consumers who purchases brand presence products are mainly loyal brand lovers and are more willing to spend.

ANALYSIS INSIGHTS

Clustering was applied to the cumulative sales data by product category. From the analysis, 6 distinctive clusters were identified. Looking at the Pseudo T-Square diagram, a sharp variance of the clusters can be observed in between of 0 – 10 scale on the number of clusters and it flattens down as the clusters increases. The sharp variance can also be observed in the Pseudo F and CCC graph.

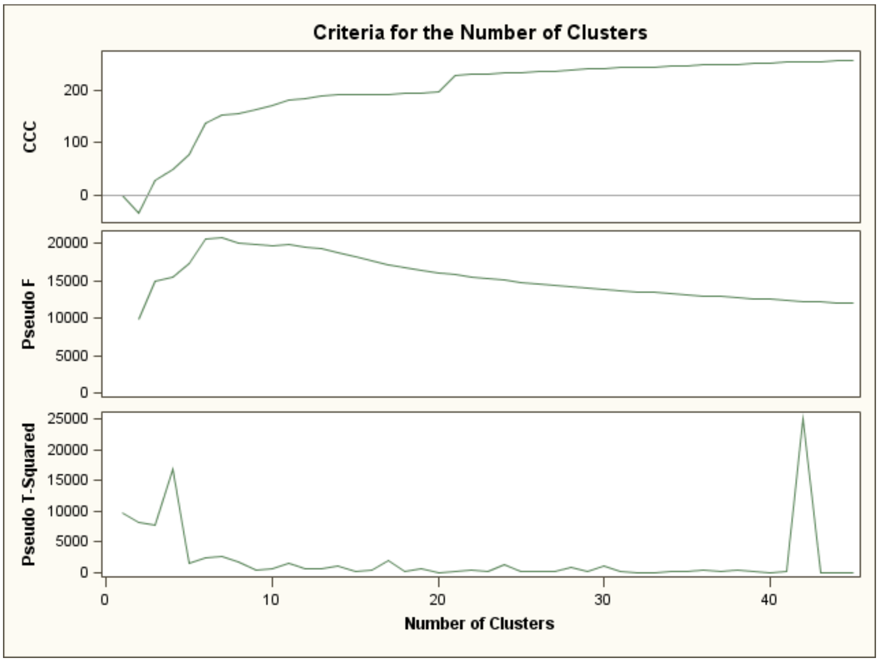


Diagram 7: Clusters

From the cluster history table below, smaller values of SPRSQ were observed as the clusters increases. This indicates further that distinctively 6 clusters were observed from the analysis. However, it was observed that there was only 1 product category that falls in one of the clusters. Hence, we can exclude the 1 unique product category and conclude that cumulative sales are classified in 5 unique clusters.

Cluster History									
NCL	Clusters Joined		FREQ	SPRSQ	RSQ	ERSQ	CCC	PSF	PST2
10	CL27	CL20	695	0.0068	0.93	0.82	172	20,000	599
9	CL32	CL15	310	0.0077	0.92	0.811	164	20,000	437
8	CL13	CL16	3230	0.0091	0.92	0.8	157	20,000	1,788
7	CL11	CL17	3707	0.0094	0.91	0.786	152	21,000	2,761
6	CL12	CL18	2463	0.0179	0.89	0.769	138	20,000	2,373
5	CL9	CL10	1005	0.0455	0.84	0.745	77.6	17,000	1,648
4	CL7	CL42	6239	0.0604	0.78	0.71	48.8	15,000	17,000
3	CL8	CL6	5693	0.0836	0.7	0.652	27.4	5,000	7,689
2	CL5	CL3	6698	0.2685	0.43	0.534	-34	9,757	8,219
1	CL4	CL2	12937	0.43	0	0	0	.	9,757

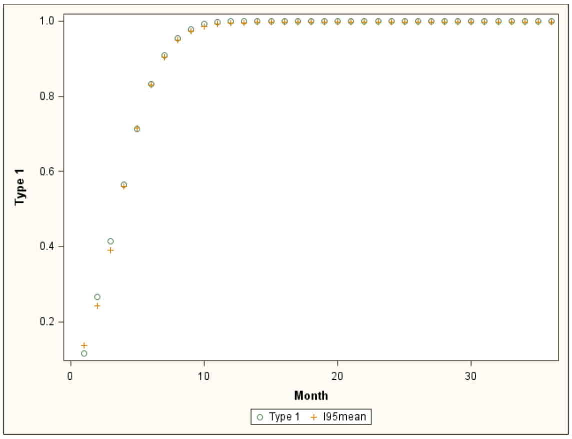
Table 1: Cluster History

GROWTH CURVE ANALYSIS

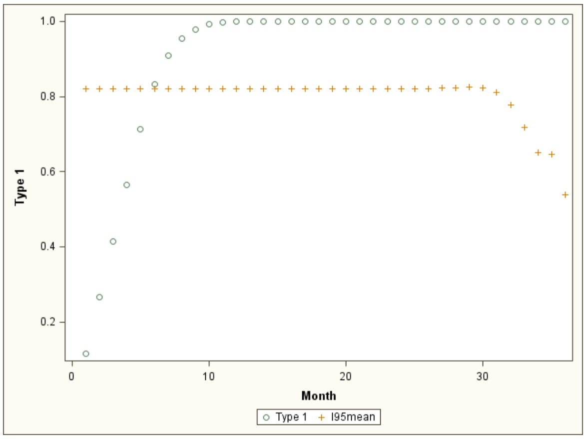
Further to the cluster analysis, four different group curves as mentioned in the earlier section were applied to the 5 different clusters to identify the appropriate growth curve to be implemented. In cluster 1, Logistic growth curve was first applied to cumulative sales and we can see that a close match was obtained.

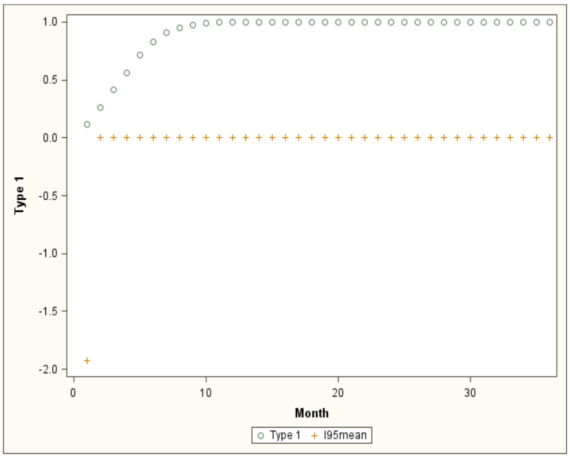
Mitscherlich curve was the second growth curve that was applied to the cumulative sales curve of the cluster. It can be seen that the Mitscherlich function was not an appropriate function to be applied to cumulative sales data as the Mitscherlich curve did not fit with the actual sales curve. The same mismatch of sales curve and the growth curves for Gompertz and Allometric were observed as well.

Logistic Curve

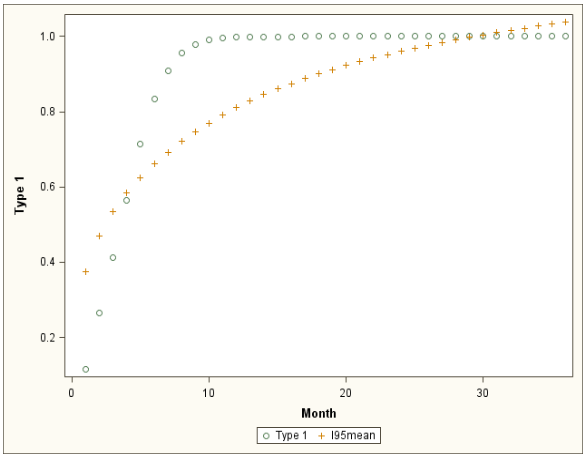


Mitscherlich Curve





Gompertz

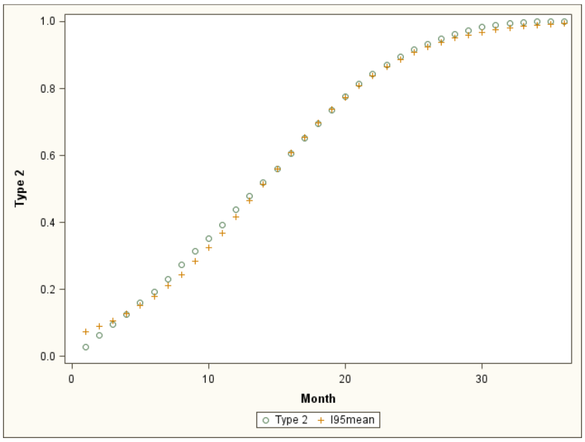


Allometric

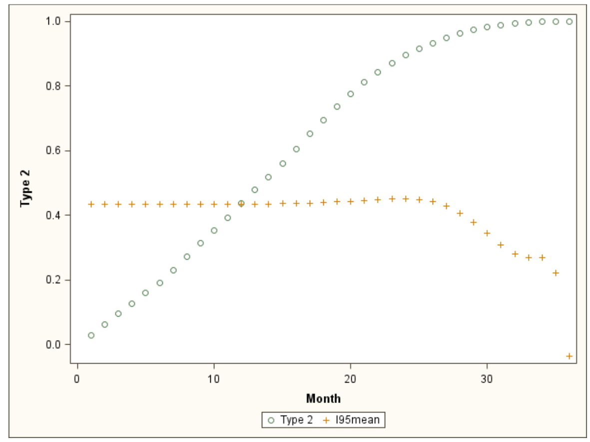
Diagram 8: Cluster 1 Growth Models

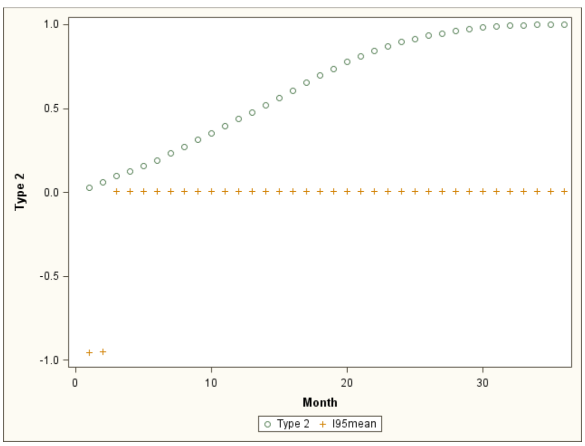
The same four growth curves were then applied to cluster 2 as shown in Diagram 9 and it was observed that a close match can be seen for the Logistic growth curve with the actual cumulative sales of the cluster. Similar to cluster 1, the Mitscherlich and Gompertz growth curves did not match with the actual sales indicating that the 2 curves are not appropriate growth models to be However, a much closer match was also observed for the Allometric curve with the cumulative sales for cluster 2.

Logistic Curve

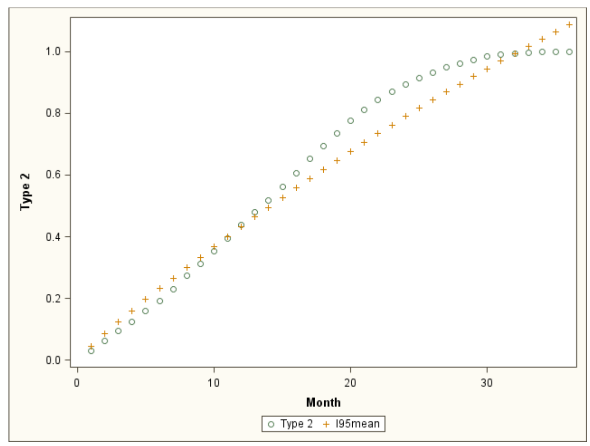


Mitscherlich Curve





Gompertz

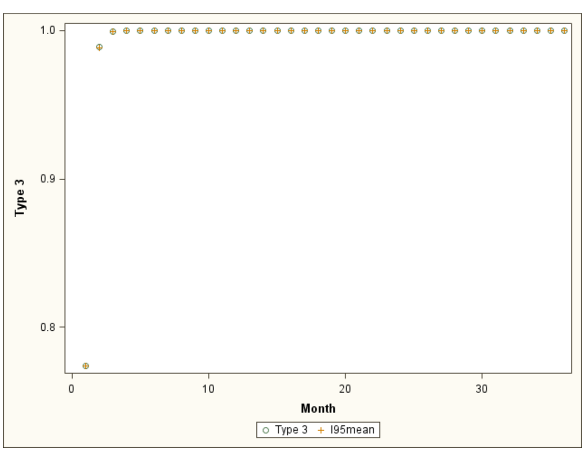


Allometric

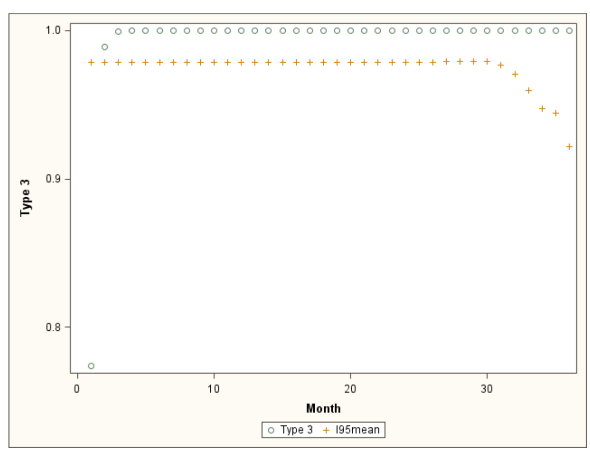
Diagram 9: Cluster 2 Growth Models

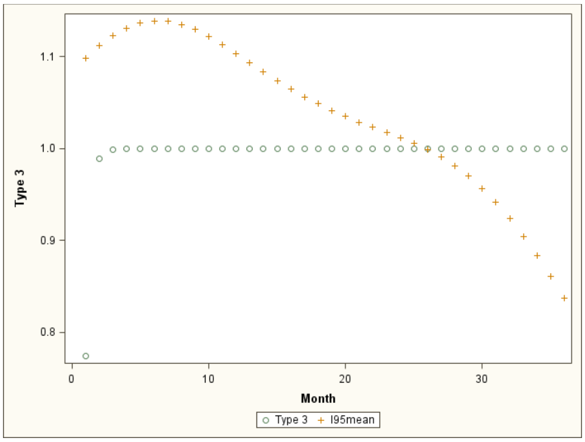
The similar curves were then applied to the remaining clusters and as seen in the following diagrams, consistent results reflect that Logistic growth curve is the most appropriate growth model that can be applied to the five clusters.

Logistic Curve

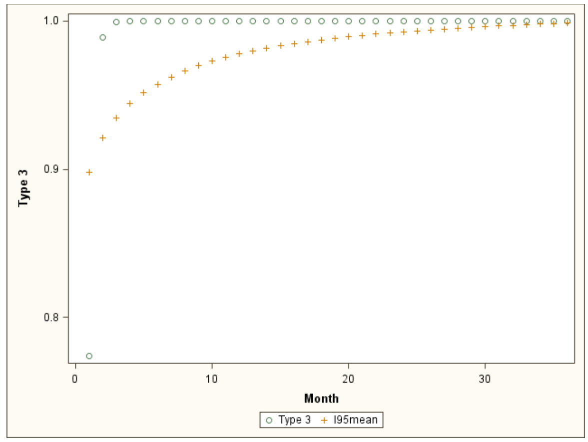


Mitscherlich Curve





Gompertz

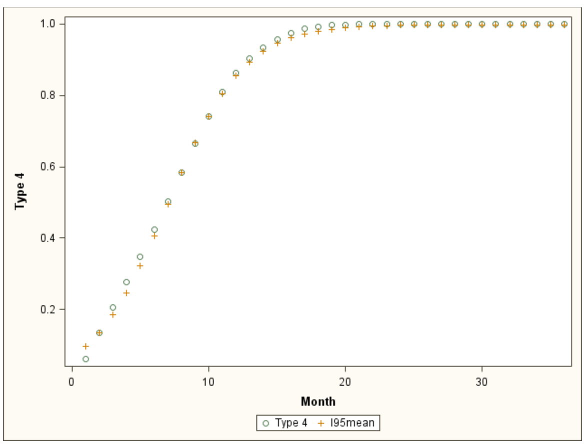


Allometric

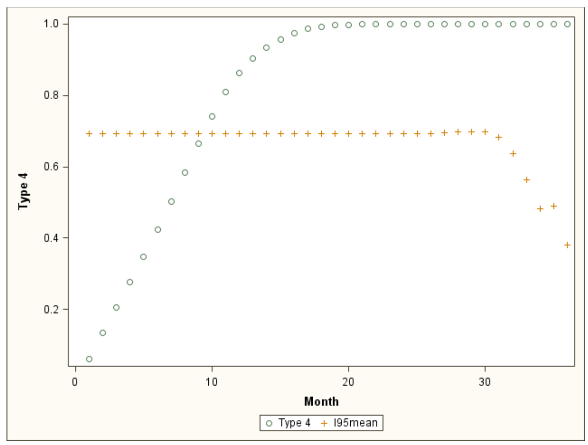
Diagram 10: Cluster 3 Growth Models

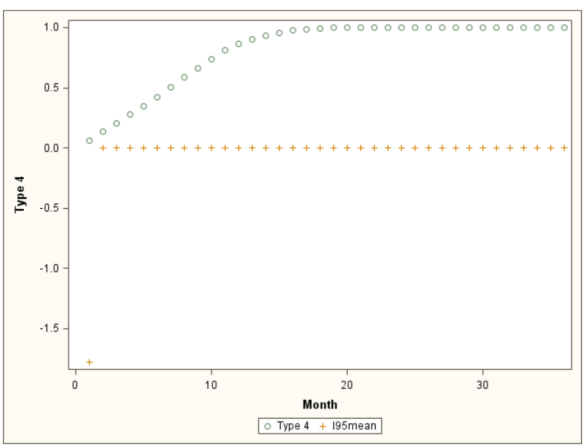
As seen in Diagram 11, the same results were reflected where a close fit was observed with the sales curve and the Logistic growth curve.

Logistic Curve

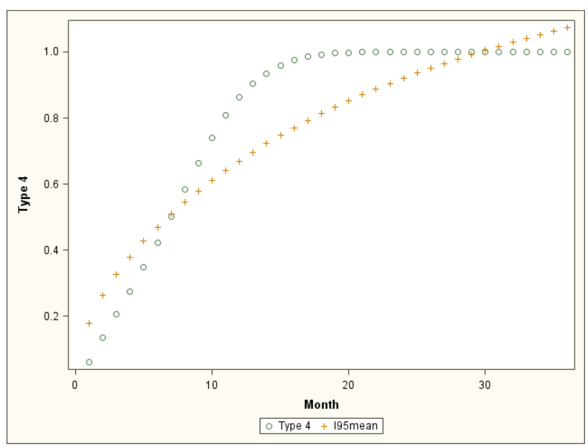


Mitscherlich Curve





Gompertz

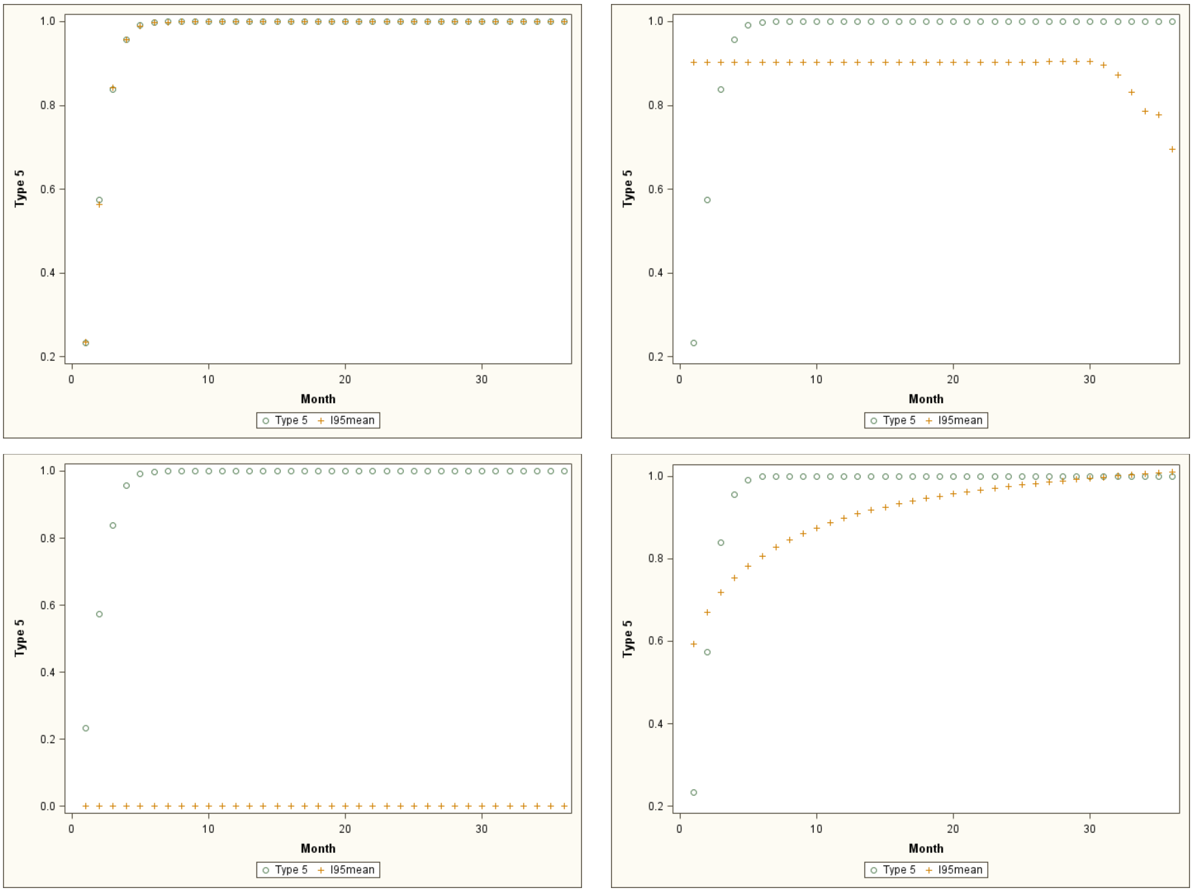


Allometric

Diagram 11: Cluster 4 Growth Models

Similarly to the other clusters, the same results were reflected for cluster 5 as shown in Diagram 12 below.

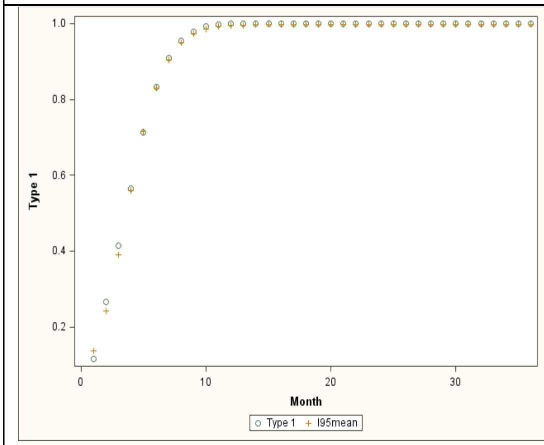
Logistic Curve **Mitscherlich Curve**

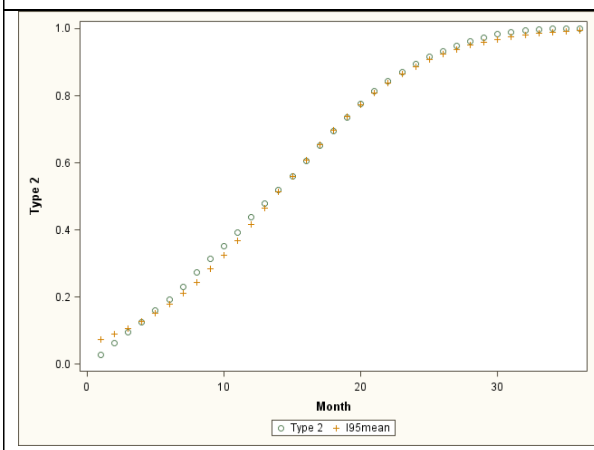


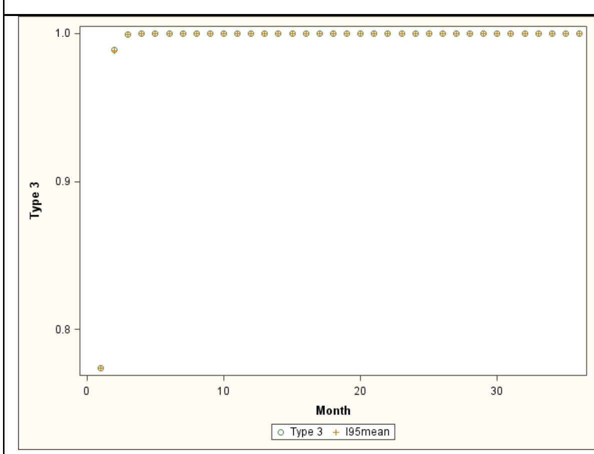
Gompertz **Allometric**

Diagram 12: Cluster 4 Growth Models

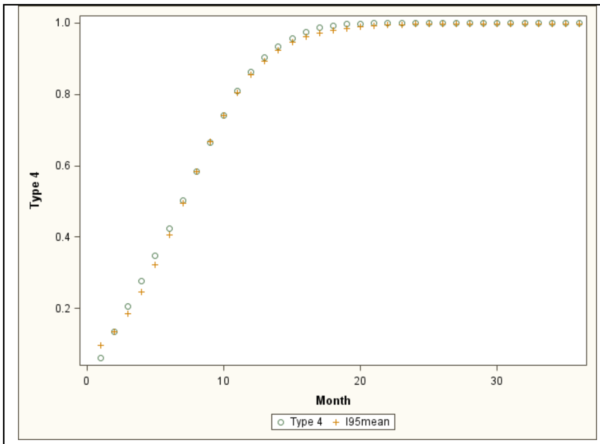
As based on the results shown in Diagram 8 – 12, a consistent result reflects that Logistic growth curve is the most suitable growth curve that can be applied for all the 5 clusters. With that, we can conclude that Logistic growth curve is the most suitable growth curve to be implemented to the cumulative sales as shown in Diagram 13 below.

<table>
<tr><th>Sales Growth Model</th><th>Equation</th></tr>
<tr><td>


Cluster 1
</td><td>f_{log} := 1.0005/1+11.7306exp(-1*1(t))</td></tr>
<tr><td>


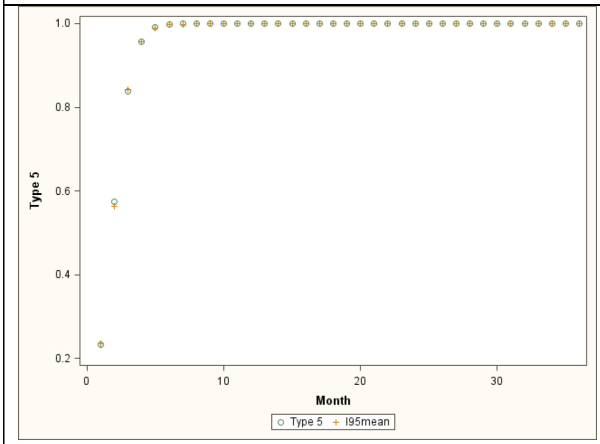
Cluster 2
</td><td>f_{log}:= -0.6931/(1+(0.8951exp(-1.000*t))</td></tr>
<tr><td>


Cluster 3
</td><td>f_{log}:= -0.1724/(1+(0.9980exp(-1.000*t))</td></tr>
</table>



Cluster 4

f_{log}:= -0.3928/(1+(0.9144exp(-1.000*t))



Cluster 5

f_{log}:= -0.1666/(1+(0.9349exp(-1.000*t))

Diagram 13: Logistic Function

IMPLEMENTATION

Having selected the appropriate growth model based on the findings above, Logistic growth curve will be applied to the cumulative sales of new watch models that will be release to the market in the coming months to determine the accuracy of the model. Further to that, analysis of the new watches will be monitored for three months.

Firstly, actual sales of the new models will be plotted against the Logistic growth curve. To ascertain that the Logistics growth curve is applicable to the new models, we need to ensure that the gap between the actual sales and Logistics growth curve is a close match, similar to that was observed in the above sections.

Analysis Of Maximum Likelihood Parameter Estimates

DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
1	-0.1667	0.0039	-0.1743	-0.1591	1844.28	<.0001
1	1.5039	0.1504	1.2092	1.7987	100.00	<.0001
1	1.7314	0.1539	1.4297	2.0331	126.50	<.0001
1	1.2261	0.1523	0.9275	1.5246	64.79	<.0001
1	0.1260	0.1794	-0.2256	0.4776	0.49	0.4823

Analysis Of Maximum Likelihood Parameter Estimates						
DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
1	0.5048	0.1751	0.1617	0.8480	8.31	0.0039
1	0.6374	0.1543	0.3349	0.9399	17.06	<.0001
1	1.4998	0.1503	1.2051	1.7945	99.51	<.0001
1	1.8059	0.1474	1.5171	2.0947	150.17	<.0001
1	1.4287	0.1563	1.1224	1.7351	83.54	<.0001
1	0.9443	0.1538	0.6428	1.2458	37.69	<.0001
1	-0.0724	0.1834	-0.4318	0.2871	0.16	0.6932
1	1.2402	0.1905	0.8668	1.6136	42.38	<.0001
0	0.0000	0.0000	0.0000	0.0000	.	.
1	0.9533	0.0167	0.9211	0.9867		

Table 2: Predictive Analysis Results

As observed on the analysis results shown in table 2, a ChiSq result of <0.001 indicates by applying the Logistic growth curve to the actual sales data, a close math of the predictive model can be identified as comparing to the actual sales. This is further illustrated in the diagrams following below.

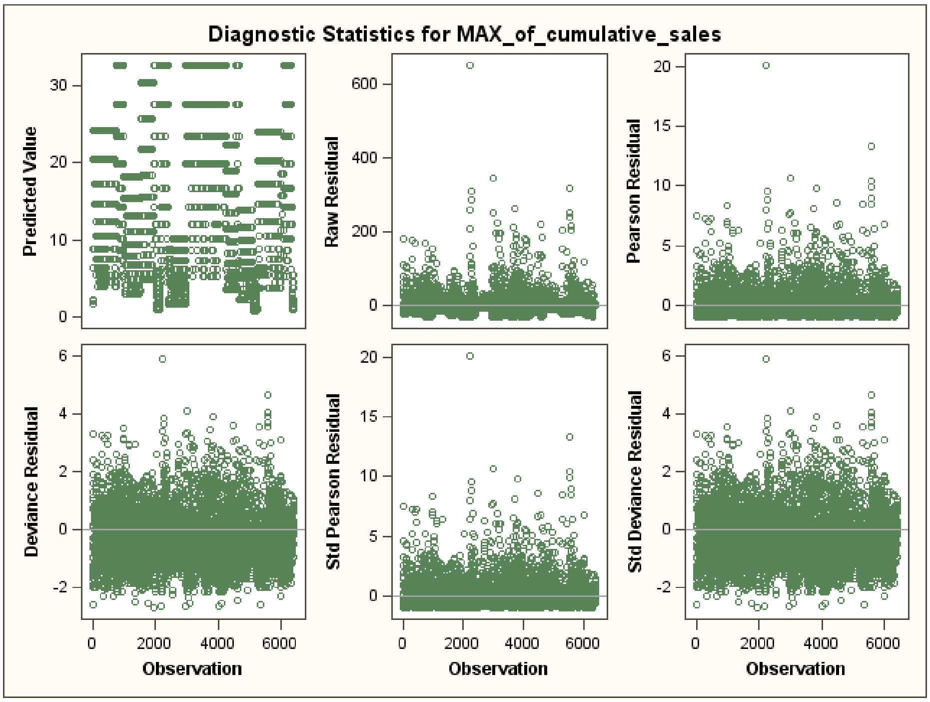


Diagram 14: Diagnostic Statistics

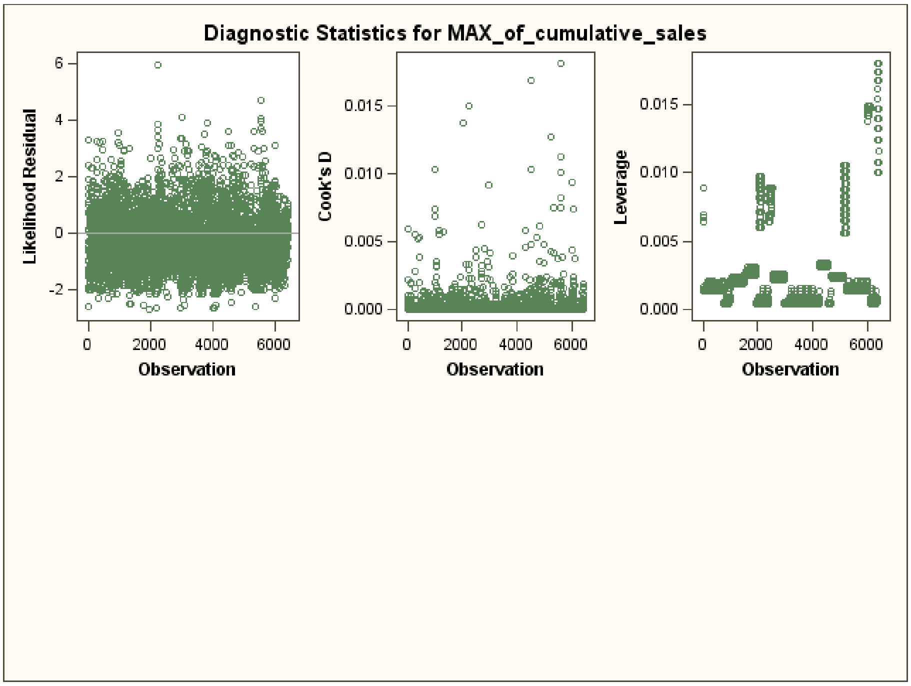


Diagram 14i: Diagnostic Statistics

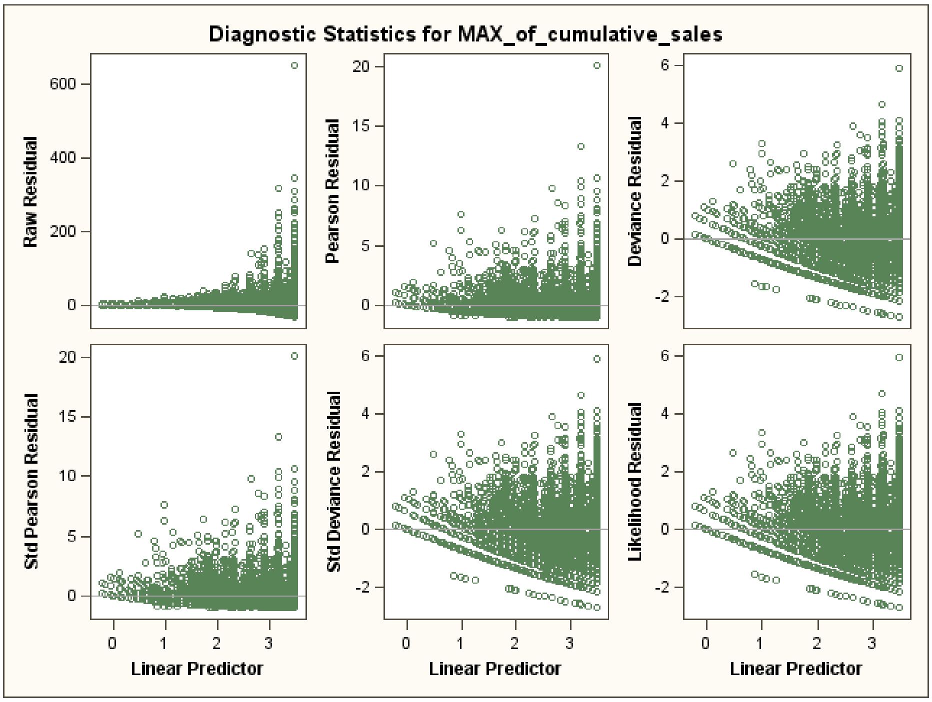


Diagram 14ii: Diagnostic Statistics

It is also important to note that the growth curve needs to be monitored on a regular basis as the sales dynamics of various watch models will change as time goes by.

CONCLUSION

In conclusion, with the appropriate growth model being implemented, it can help the retailers to better manage their inventory as they will be able to forecast more accurate sales demand whenever there are new releases in the market. Furthermore, retailers will have a clear direction on when to trigger new sales so as to ensure stable revenue for the company. Having a more consistent sales growth, it will help retailers to understand the customer demands and introduce more suitable watch models into the market.

With the use of analytics, retailers are able to make fact-based decision instead of using their gut-feel assumptions. Past consumer behaviour data can be used to develop models to predict future customer behaviour. Having predicted future customer behaviour, it will create a more personalized relationship between the retailers and consumers thus bringing the retailers to a new level of customer centricity.

REFERENCES

Pedro Pereira Rodrigues. 2005. Hierarchical Clustering of Time Series Data Streams. Pg. 31 – 38

Michael Falk, Frank Marohn, Rene Michel, Daniel Hofmann, Maria Macke, Christoph Spachmann, Stefan Englert, 2006, A First Course on Time Series Analysis – Examples with SAS, Chair of Statistics, University of Wurzburg

NICHOLAS P. JEWELL & ALAN HUBBARD, January 18 2006, Analysis of Longitudinal Studies in Epidemiology, Chapter 4 Modelling Counts – The Poisson and Negative Binomial Regression

Thomas H. Davenport, 2009, Realizing the Potential of Retail Analytics, Babson Executive Education

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Name: Cally Ong Yeru
Address: Blk 273B Bishan Street #09-108
City, State ZIP: Singapore, 572273
Work Phone: +65 9843 2015
E-mail: yeru.ong.2012@mitb.smu.edu.sg