

## Assignment #1 Case Study: Exotic Apparels

### ***Background:***

Exotic Apparels boasted of an impressive chain of retail outlets selling premium apparel in almost every metro in Europe. The company started in Holland in 1885 and grew in to a continent-wide business by 1950 primarily focusing on men's apparels. Between 1970 and 2000 the company expanded into high-end women's wear and premium footwear. Business expanded rapidly in the last 20 years with many new stores and customers being added to the company's database at an impressive rate. Worldwide revenues stood at 152 million dollars in 2003. The five-year annual revenue growth rate was 6.5 %. However, it was increasingly unclear to Mathew Tufte, the Managing Director of Exotic, which of their customers would turn out profitable in the long run and which of them would not.

He wondered if by observing past purchase behavior there was a way to predict the future profitability of a customer. He commissioned a research team to collate and access customer data from all the retail stores in Holland to estimate the Lifetime Value of each individual customer. He believed that by doing this it might be possible to identify superior customers (those customers who were profitable) in advance helping the company provide better service and targeted products to them. Mathew was also keen to understand which of his stores was better positioned to deliver profits in the coming three years, as he needed to plan for investments in various stores and the renewal of lease/rental agreements. Mathew had a hunch that a lot of his old time customers were not exactly the most profitable in the long run. But, he needed proof that this was indeed the case. He was also interested in checking out if the RFM approach adopted by the industry to select top customers was actually applicable to his company, or if another method fit more closely to his company's needs.

### ***Identification of research areas:***

The research team identified the following areas of analysis:

1. Determine the Customer Lifetime Value (CLV) of each customer.
2. Analyze the profiles of customers segmented based on CLV.

3. Determine the profitability of each store.
4. Define and measure the extent of loyalty for each customer based on traditional metrics (Regularity of purchase, Frequency of purchase, RFM score and Tenure of the customer).
5. Determine if customers selected on the basis of traditional metrics are the right choice from the view of long-term profitability. Check to see if CLV is a better indicator of long-term profitability than the traditional metrics.
6. Determine what factors drive profitability and compare them with the factors that drive loyalty.

***Data requirements:***

The research team then determined the data required to conduct the analysis by breaking down the components of the CLV framework. During discussions with the Exotic management in Holland, it was decided that the time period for the CLV computation would be restricted to a three-year period, since this was the typical planning horizon that they followed. The research team examined the rich customer information that Exotic had been collecting over several years and requested the management to provide data pertaining to their customers along the following dimensions:

1. Unique number that identified a customer in the database across all the transaction that he/she makes.
2. Demographics describing each customer (Age, gender, occupation, income, marital status, number of children, location, distance of residence from a retail store, etc.)
3. The shopping behavior exhibited by customers while buying Exotic products in terms of:
  - a. Date of each transaction (including product returns)
  - b. Value of each transaction (including product returns)
  - c. Product category purchased in each transaction.
  - d. Channels of shopping (e.g. mail order, retail store, web sales, factory outlet)

4. Lifestyle variables such as:
  - a. Membership of clubs
  - b. Nature of other shopping venues frequented by the customer

***Metrics computed:***

The research team proceeded to compute the following metrics from the available data as a first step towards determining CLV and profiling its customers:

1. Frequency of purchases made by every customer based on each customer's purchase history.
2. Total revenue contributed by each customer through each channel of purchase until the current date.
3. Revenue contributed by each customer by product category.
4. Average dollar spent per transaction by each customer.
5. Total tenure of each customer (the period from the first purchase until today).
6. Recency of the last purchase made (the period from the last purchase until today).
7. Extent of cross buying (the total number of different product categories purchased by each customer).
8. The total number and value of returns made by a customer.
9. Primary channel of purchase (retail, mail order, outlet mall, etc.).
10. Seasonality of buying behavior of each customer, if any, and classify customers as Holiday shopper, fall shopper, spring shopper etc.
11. Cost of marketing activities directed at each customer.
12. Expenses incurred at the store level based on each transaction.

***Developing the CLV model:***

The two components of the CLV model that needed to be developed were:

1. **The purchase frequency model:** This was required to predict the frequency with which a customer will continue to buy in the future (average number of purchases made every year for the next three years).
2. **The gross contribution model:** This was required to predict the gross contribution that can be expected from the customer in the future (the next three years).

The purchase frequency for any customer and the gross contribution provided by the customer depend on the purchase history of the customer, the demographic and lifestyle profile of the customer, and also the marketing contact efforts made by the company. The research team had already obtained and organized the data for the range of metrics that they a priori felt would help them model both these components of CLV. A sample of customers was chosen and a multi-spell hazard model<sup>1</sup> was developed to understand which metrics would help predict the purchase frequency of each customer most closely. Similarly a multivariate regression model was developed to understand the variables that would predict the gross contribution from each customer. Several variables were tested for their explanatory power and only those variables that strengthened the two models were retained (See Appendix I). The models were built using 75 % of the customers sampled. The results of the model were applied on the remaining 25 % of the customers to test for predictive accuracy. It was examined if the purchase frequency and gross contributions predicted by the models were close enough to actual values observed for this balance 25%. For this purpose customers were divided in to deciles based on the predicted values. The observed values were then plotted for these deciles alongside the predicted values (See Appendix II). A comparison seemed to indicate that the models seemed to be fairly accurate in their predictive capability.

### ***Measurement of CLV and customer profiling:***

---

<sup>1</sup> A multi-spell hazard model takes into account multiple purchase occasions in the future instead of just one.

The CLV computation involved merging the two components. By predicting the purchase frequency and the expected gross margin followed by overlaying these predictions, we arrive at the margins each customer is expected to provide for the three-year time horizon. The next step was to subtract the corresponding marketing expenses and store expenses over the three-year time period to arrive at the accumulated margin. All this while, it is important to discount all margins and costs incurred at various time points down to the current period using the company's cost of capital as the discount rate. The net present value of each customer for the next three-year period was now available. This provided the final estimate of CLV.

Customers were then organized in to deciles based on their estimated CLV. It was important to understand the characteristics of customers who fell in to the High CLV deciles and the Low CLV deciles in order to distinguish between each of the customers. Customer characteristics that changed discernibly across CLV segments were then reported (See Appendix III).

***Drivers of CLV:***

A regression was then run with the CLV computed for every customer as the dependent variable. Several variables for which customer-level data was available were included as independent variables. The top drivers that emerged as good indicators of CLV were:

1. Lifetime revenue accrued from web sales and factory outlets
2. Tenure of the customer
3. Dollars spent on high end women's apparel
4. Degree of cross buying
5. Dollars spent on Home Furnishings
6. Returns made during the lifetime

While the first five independent variables impacted CLV in a positive manner (i.e. an increase in the independent variable led to an increase in CLV), the sixth variable exhibited an inverted U-shaped relationship (i.e. initially an increase in Returns led to an

increase in CLV, but after a threshold level, an increase in Returns led to a decrease in CLV) (See Appendix IV).

***Store Profitability based on CLV:***

While customers could shop from multiple stores, and their CLV's were based on transactions across various stores and channels, Mathew was interested in determining the profitability of each store. The research analysts computed store profits by summing the CLV contributions of all customers who were expected to shop at a particular store and subtracting the NPV of rental expenses expected for each individual store.

Interestingly, this analysis indicated that some stores located at premium locations were indeed low on profits.

***Traditional Loyalty metrics:***

Loyalty means different things to different people. However marketers have commonly used some standard measures when measuring loyalty. Three such traditional measures were identified and reported by the analysts:

1. How reliable is the customer? Does the customer purchase consistently in terms of the time gap between purchases? This was measured in terms of the variance in the inter-purchase time interval of a store customer.
2. What is the RFM score of the customer? This was computed as a weighted sum of the Recency (40%), Frequency (40%) and Monetary value (20%) provided by a customer in the past.
3. How long has the customer been transacting with the company? Hence what is the duration of association of the customer with the company?

Customers sorted in to deciles based on CLV were examined for these loyalty measures (See Appendix V). The correlations between CLV and these measures were not high at all suggesting that any of these measures taken individually do not have a direct link with CLV. The 'r' values, which determine the level of correlation between two or more variables (close to 0 is no correlation, close to -1 is perfect negative correlation, and close to 1 is perfect positive correlation), can be found in the table below:

*Correlation between CLV and traditional loyalty measures*

	Reliability/Consistency	RFM	Duration
CLV	r= 0.05	r= 0.25	r=0.45
N	3081	4225	4225

*Strategic Initiatives:*

Based on the analysis, the research team suggested that Mathew take steps to maximize CLV. They recommended the following:

1. Target high spenders in web sales and factory outlets
2. Deploy marketing tactics to encourage cross-buying
3. Identify customers exhibiting high incidence of net returns and devise strategies to discourage that behavior
4. Encourage customers who spend in high end women's apparel Target customers exhibiting high duration.

*Questions to answer in this case analysis:*

1. By focusing only on profitable customers based on CLV, is the company doing the right thing from a long-term business point of view? What are the other decisions that the company can take based on CLV?
2. Comment on the strategy recommended by the research team. Are there any strategies that you would add or remove from this list?
3. How different would the recommended strategy be if traditional loyalty metrics were used instead of CLV?
4. It seems intuitive that the first five variables listed have a positive correlation with CLV. Why do you think that the amount of product returns has an inverted U-shaped relationship with CLV?

## **Appendix I**

### ***1. Purchase Frequency Model***

The predictions are based on:

- The timing of customer’s past transactions in retail stores and other channels.
- The categories from which a customer purchases in all channels
- The profits and revenues provided by a customer in each transaction.
- The stores in which a customer transacts.

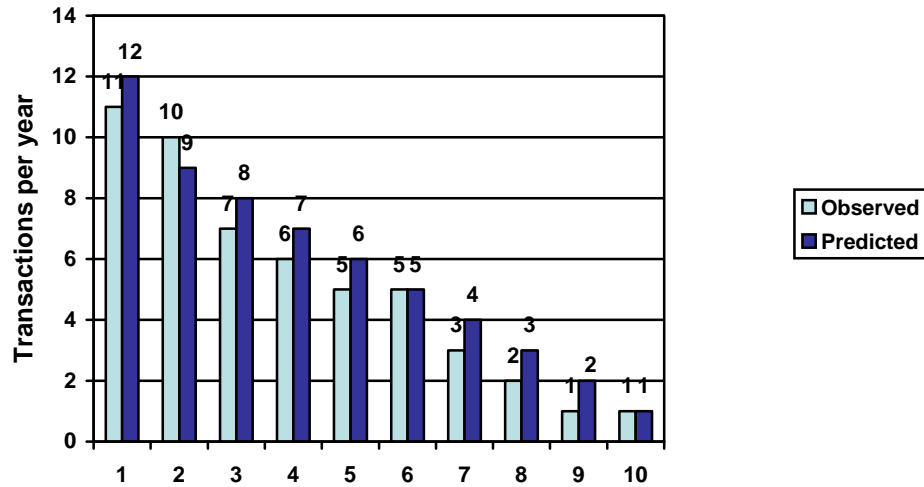
### ***2. Gross Contribution model***

The predictions are based on:

- Past Gross Contribution Margin
- The categories from which a customer purchases in all channels
- The stores in which a customer transacts

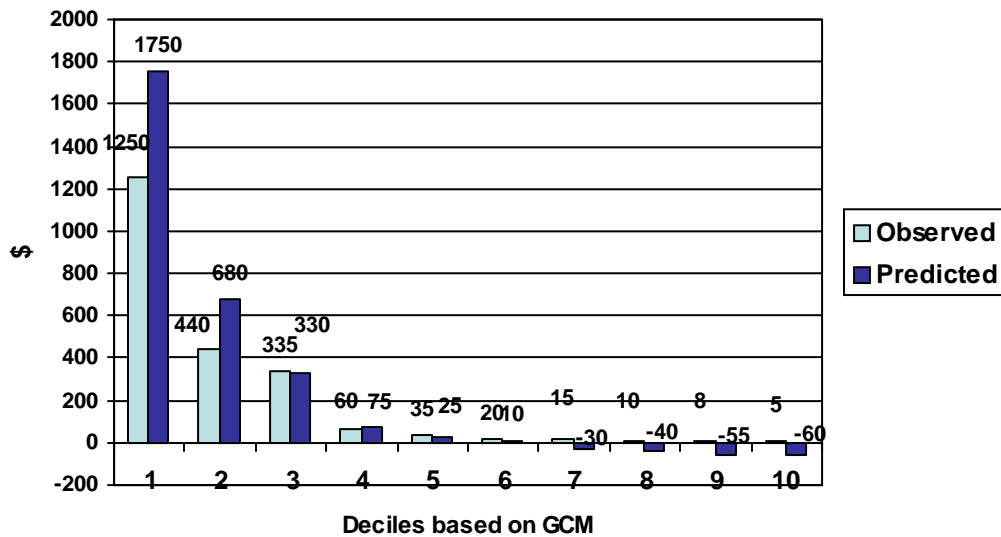
## Appendix II

### Predicted versus Observed Purchase Frequency



Deciles based on Predicted Frequency

### Predicted versus Observed GCM



### Appendix III

<b>High CLV customer profile</b>
Age Range: 35-54 years
Gender: Female
Estimated Income: \$ 80,000
Catalog Shopper
High frequency shopper in premium department stores
Marital Status: Married
Has children
High proximity to company store

<b>Low CLV customer profile</b>
Estimated Income: < \$45,000
Age Range: 25- 35 years
Marital Status: Single
Low proximity to company store
Gender: Male
Does not shop across multiple channels

## **Appendix IV**

### **Top Drivers of CLV**

1. Lifetime Revenue from web sales and factory outlets (+)
2. Duration (+)
3. \$ Spent in High end Women's Apparel (+)
4. Cross-Buying (+)
5. Lifetime Returns ( $\cap$ )

### **Top Drivers of Reliability/ Consistency of Purchase**

1. Total Lifetime Revenue (+)
2. Estimated Income (+)
3. Overall Cross-Buying Including web sales and factory outlets (+)
4. Average \$ spent per transaction (+)
5. Indicator that the customer is predominantly a Sale Shopper (+)

### **Top Drivers of RFM**

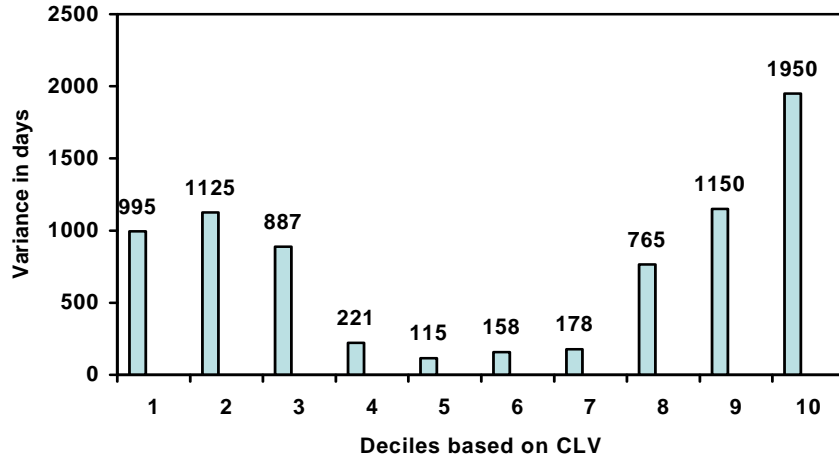
1. Estimated Income (+)
2. Indicator that the customer purchases predominantly from retail store channel (+)
3. Marital Status (+)
4. Fall Customer (-)
5. Total returns in current year (-)

### **Top Drivers of Duration**

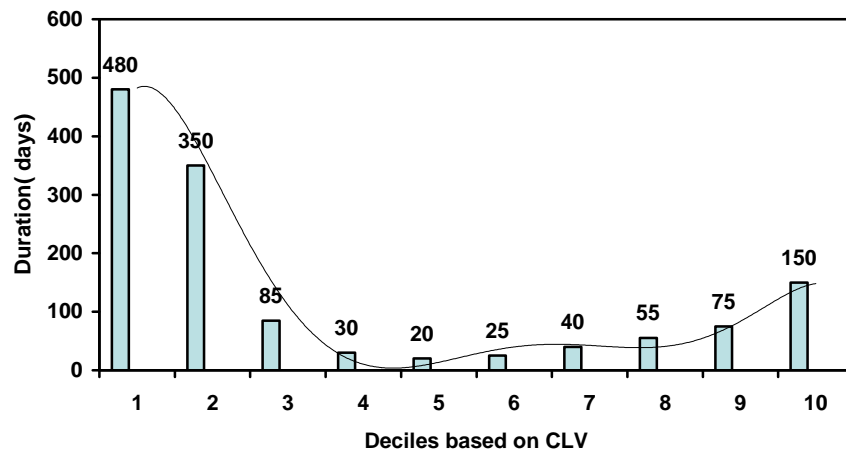
1. Cross-Buying (+)
2. Fall Customer (+)
3. Frequency of Visits (+)
4. Lifetime Transactions (+)

## Appendix V

### Reliability/Consistency of Purchase



### Duration



### RFM score

