BUS5PA Tutorial 7
Customer Retention and Churn Analysis

Objectives of the tutorial

1) Understand the key concepts customer churn/attrition and retention
2) Discuss the main features of attrition and retention with a series of questions/answers
3) Use SAS Enterprise Miner to build models of customer churn.
4) Use a scoring data set to score new customers
5) Carry out a SAS Enterprise Miner based exploration of churn results

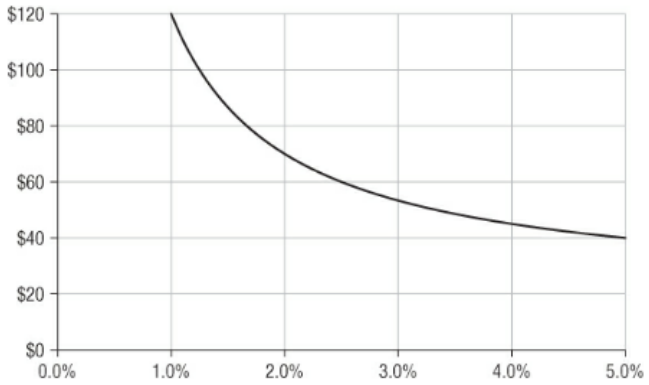
Background

Customer attrition (churn) is an important issue for any company, and it is especially important in mature industries where the initial period of exponential growth has been left behind. Therefore attrition (and retention) is a major application of data mining where predictive analytics plays a key role.

One of the first challenges in modeling attrition is deciding what it is and recognizing when it has occurred. This is much harder in some industries – for e.g.: businesses that deal in anonymous cash transactions. Attrition is a bit easier to spot when a monthly billing relationship exists, as with credit cards. But even in such cases a customer may stop using the credit card, but not cancel it. Attrition is easiest to define in subscription-based businesses, and partly for that reason, attrition modeling is most popular in these businesses (mobile phone service providers, insurance companies, cable companies, financial services companies, Internet service providers, newspapers, magazines, and some retailers all share a subscription model where customers have a formal, contractual relationship that must be explicitly ended).

Why is customer retention so important?

Lost customers must be replaced by new customers, and new customers are expensive to acquire. Often, new customers generate less revenue in the near term than established customers.



As the market becomes saturated and the response rate (X axis) to acquisition campaigns goes down, the cost (Y axis) of acquiring new customers goes up. **At some point, it makes sense to spend that money to hold on to existing customers rather than to attract new ones.**

Retention campaigns can be effective, but also expensive. A mobile phone company might offer an expensive new phone to customers who renew a contract. A credit card company might lower the interest rate. The problem with these offers is that any customer who is made the offer will accept it. Many of the people accepting the offer would have remained customers anyway. **The motivation for building attrition models is to figure out who is most at risk for attrition so as to make the retention offers to high-value customers who might leave without the extra incentive.**

Business Case

A telecommunications company that offers voice, messaging, and data services has noticed a high rate of churn among its high-revenue customers in recent years. A customer retention campaign was developed and the marketing managers need to identify the customers who have a high probability of churning. These targeted customers are then placed in the customer retention program.

The probability of churning is also known as a ***churn propensity score***. Customers with high scores are more likely to churn and are thus candidates for the customer retention program.

Possible predictor variables are:

Outstanding bill value
Outstanding balance period
Number of calls
Call duration (international, local, national calls)
Period as customer
Total dropped calls
Total failed calls

Customer demographics such as age, education, and gender can also be useful predictor variables.

After a predictive model is developed and selected, it is used to score customers to obtain the probability of churning.

Steps to follow:

A) Create library, new project, data source and diagram

B) Manage missing values and change values

C) Variable clustering and build regression and decision tree model

D) Explore the data to understand causes of the unusual model performance

E) Refit regression and decision tree models to the modified data

F) Scoring a new data source

Homework

Answer the following questions – refer Data Mining techniques, Berry and Linoff as reference. (Chapter 2, pages 60-64 and chapter 10 pages 357 -364)

1. What are the main kinds of attrition? What is the differences and significance of these?
2. There are two main kinds of attrition models – modelling attrition as a binary outcome, estimating customer life time.
3. What is the difference and what are the benefits and applications of each type of model?
4. What is the most important facet in customer behaviour? How is this related to attrition/churn?
5. Survival analysis is a term used to using analysis and analytics techniques to understand 'when to start worrying' about customers (in a business/marketing context). Compare and contrast data volumes used for survival analysis in business applications with medical and manufacturing applications.
6. What are the key differences in using statistical techniques and data mining techniques for survival analysis? (refer to the article data mining and statistics in the additional reading material folder).
7. What are the key events captured by survival analysis? Are these related to churn analysis?
8. What is a survival curve? How could survival curves be useful in customer retention?