



# **How to Build a Marketing Database *and Use it to Get More Sales and Profits***

By Perry D. Drake

Westchester Direct Marketing Forum – April 15, 2003

# Objectives

- An Introduction to Database Marketing
- The Components of a Marketing Database
- What is Entailed in the Build of a Marketing Database?
- Building the Database In-house or Outsourcing – How to Decide?
- An Introduction to Segmentation
- CHAID Analysis

# Perry Drake's Background

## ■ Current Responsibilities:

- Vice President and General Manager, Drake Direct, New York, NY
- Associate Professor, New York University, Center for Direct Marketing, New York, NY. Currently teaching "Database Modeling and Analysis" and "Advanced Database Modeling." 1999 Recipient of the "*Outstanding Master's Faculty Award*."
- Instructor, Western Connecticut State University, Interactive Direct Marketing Certificate Program, Danbury, CT. Lectures on testing and marketing financials.

## ■ Prior Responsibilities:

- Director, Marketing Services, The Reader's Digest Association, Pleasantville, NY
- Associate Director, Magazine Marketing, The Reader's Digest Association, Pleasantville, NY

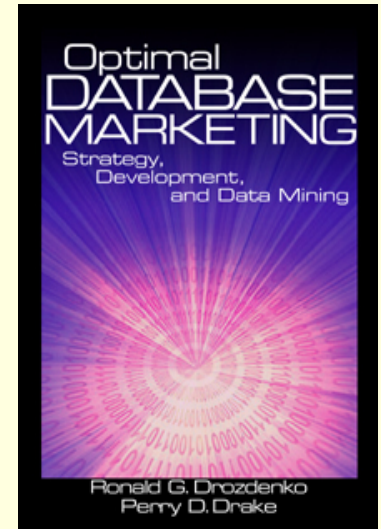
## ■ Education:

- Master of Science, Applied Statistics, University of Iowa
- Bachelor of Science, Economics, University of Missouri

# Perry Drake's Background

## ■ Publications and Presentations:

- “*Optimal Database Marketing*,” a new publication authored by Perry D. Drake and Dr. Ron Drozdenko, Sage Publications. This new industry book delves into database marketing concepts and practices covering topics such as the process to evaluate database needs and then select a database vendor, analyzing and manipulating the customer data, segmenting the customer file, response modeling, strategic reporting including lifetime value calculations, and test design and analysis.
- Jointly with other faculty from Western Connecticut State University, he collaborated with the Direct Marketing Educational Foundation to develop a model curriculum for universities pursuing the area of interactive and direct marketing.
- A complete listing of Perry's publications and industry presentations can be found at [www.DrakeDirect.com](http://www.DrakeDirect.com).



# Database Marketing - Introduction

Direct marketers have an advantage over traditional marketers in that we can accurately research and test concepts (offer, price, product/service) and gauge various consumer/business group reaction to those tests to ensure we make the best decisions possible regarding what to promote and to whom.

That is, if we:

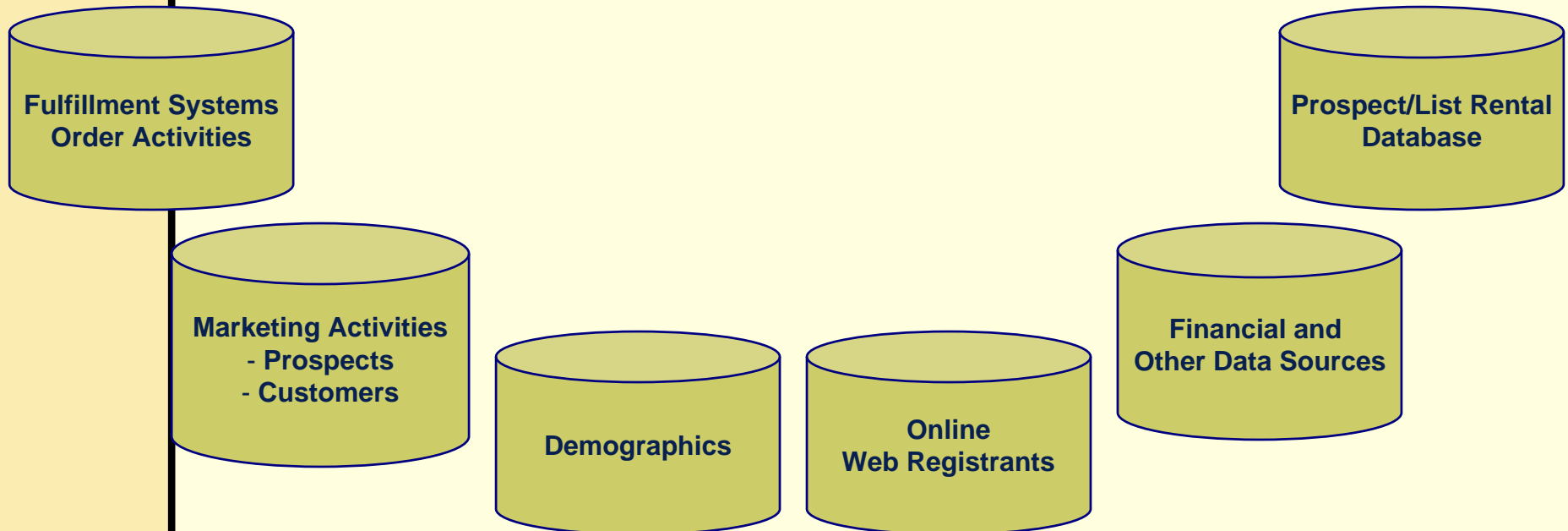
1. Have a properly built marketing database, and
2. Understand the basics of database marketing analytical applications!

# An Ideal Marketing Database will have ....

1. **Fully integrated** marketing information – marketing activity, fulfillment, demographics, order history
2. Fast, **cost-effective** implementation
3. Integration process exposes and cleans up data problems
4. Data loaded and **summarized quickly** – real-time, daily, weekly
5. Both a prospect and customer **marketing data repository**
6. **Information for everyone** – senior execs, data analysts, middle mgrs, junior level.

# A Major Issue to Overcome....

Multiple, disparate data sources...



# And the Information must be....



Data Analysts



Marketing  
Directors



Executive  
Management



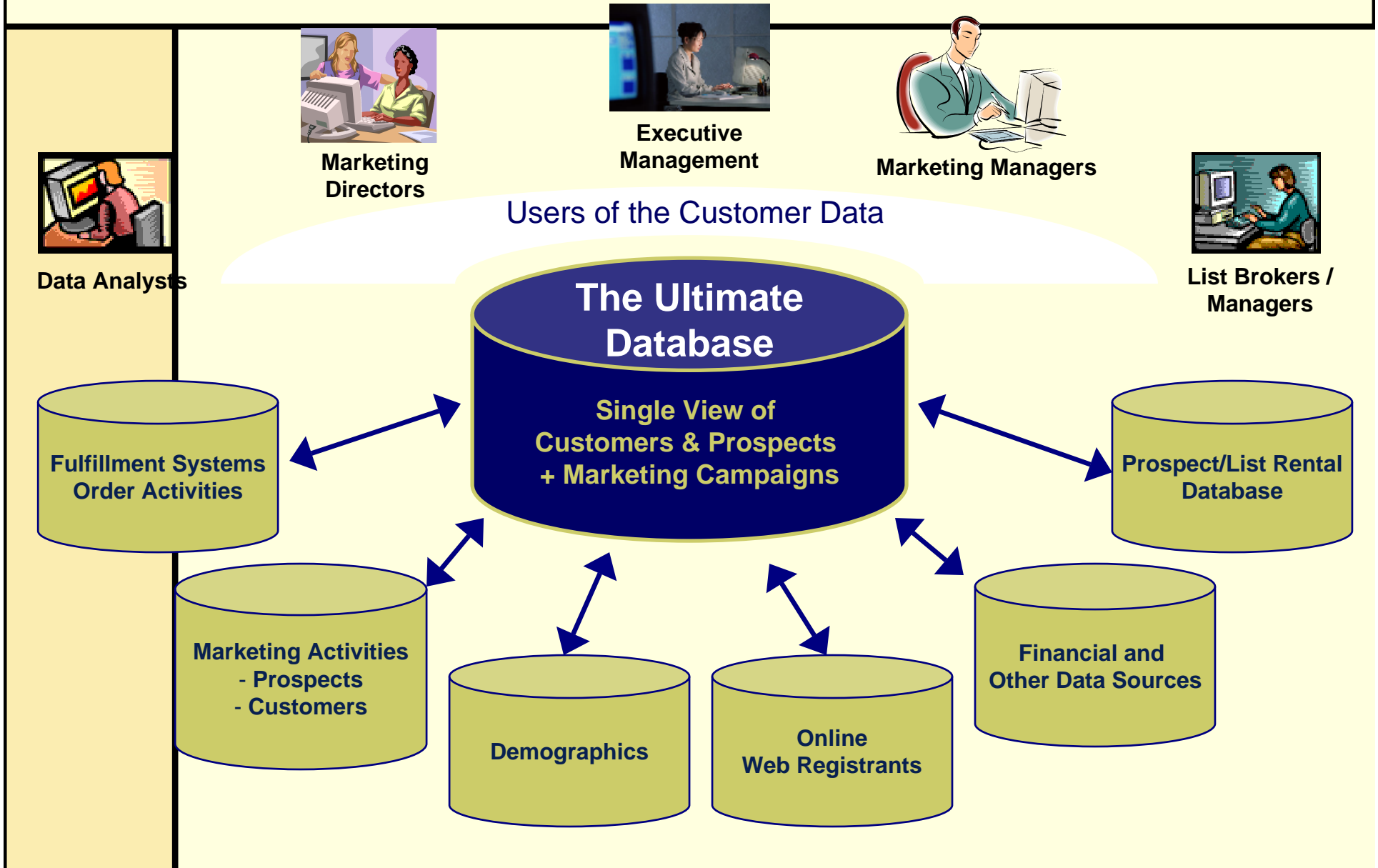
Marketing Managers



List Brokers /  
Managers

accessible by many people  
across the enterprise quickly and  
cost-effectively.

# The Complete Picture



# Next, Tools Must be Added for Accessibility



Data Analysts



Circulation Directors



Executive Management



Marketing Managers

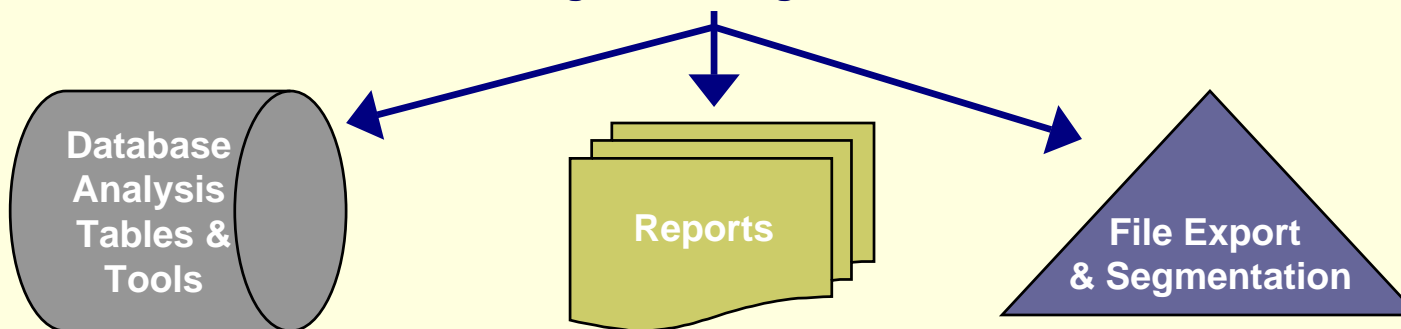


List Brokers / Managers

Users of the Customer Data



Well-Integrated & Engineered Data



# Building Database In-house or Outsource?

The build of a marketing database is quite difficult as can be seen in the prior slides.

You need to have many skills available to the organization to pull it off properly.

If you lack the technical expertise regarding computer hardware, platform and software issues you may be well advised to outsource the build.

Doing so will then free up your time to ensure the resulting database will meet user needs letting outside experts worry about technology -- not a bad thing in deed!

# Building Database In-house or Outsource?

If you do decide to build in-house, some of the technical issues you will need to determine include:

- Database Hardware – e.g., IBM Mainframe
- Operating System – e.g., Windows NT, UNIX
- Database Software – e.g., Oracle, Sybase, DB2
- Database Structure – e.g., Flat or relational
- Query Language – e.g., SQL for name extraction, SAS for analysis

*What you choose depends on your business and user needs.*

# Building Database In-house or Outsource?

If you do decide to outsource you will need to gather internal needs to ensure all requirements are fully met once the database is complete and delivered.

In particular, gather data and reports used and frequency of use for each functional area including:

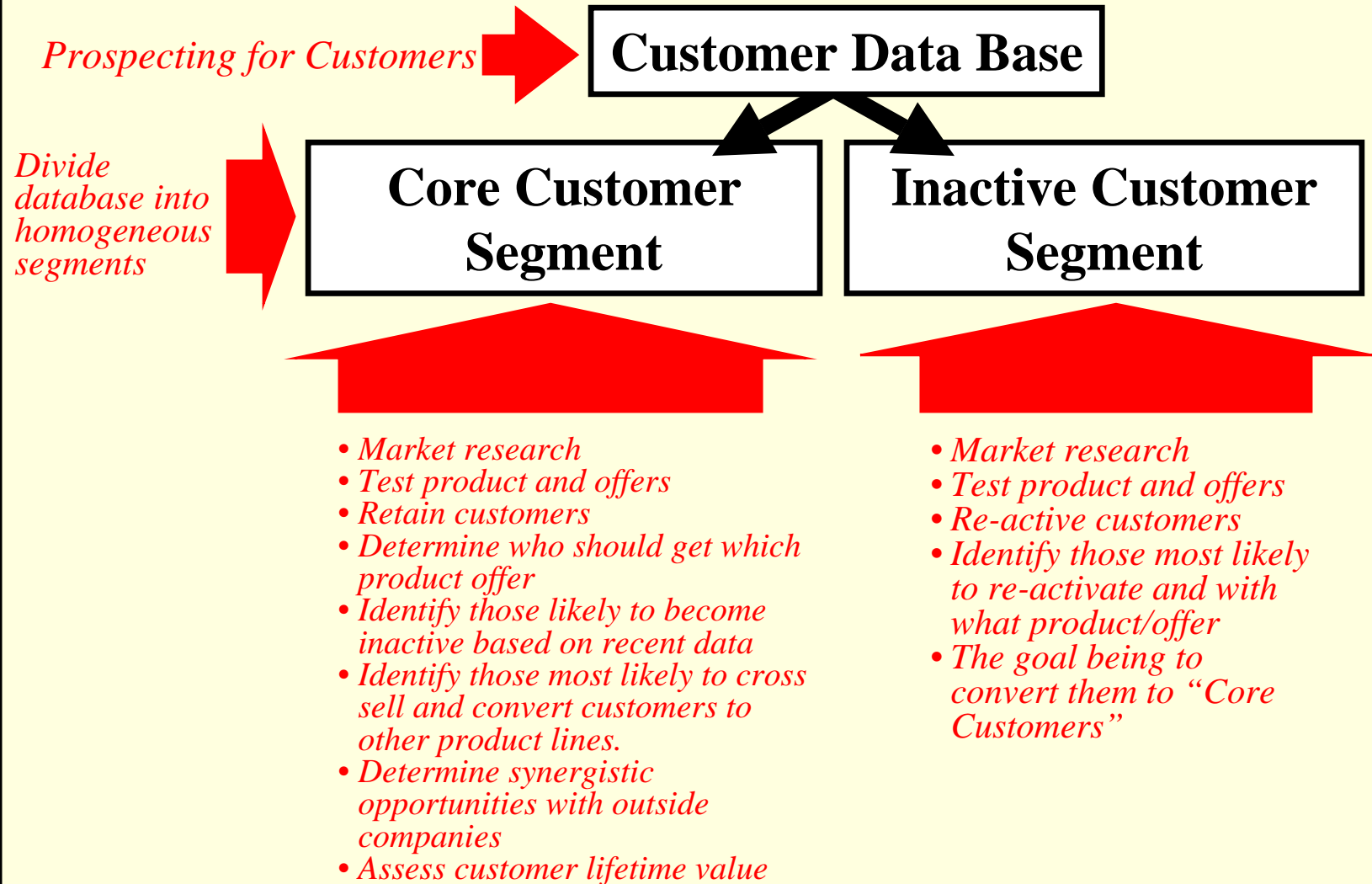
- Marketing
- Finance
- Corporate
- Customer Service
- Fulfillment
- List Marketing
- Online Marketing

# Building Database In-house or Outsource?

Once obtained you will then need to formalize the needs into detailed specifications covering such things as:

- Size in terms of number of records
- Data feeds
- How the feeds relate to each other
- Data tables
- Timing of data flow
- File maintenance requirements and updating rules

# Database Marketing - Activities



# Database Marketing - Activities

Given our time constraints here this morning, we will only focus on one of these key database activities -- namely segmentation.

# Segmentation - Introduction

The underlying premise for segmentation of your database is that not all customers residing on the database are alike and, therefore, should not be treated alike.

Segmentation is the process of dividing the total market into groups of people with similar needs and desires based on their characteristics and past purchase behavior with the ultimate objective being to develop a more personal and responsive relationship with customers.

# Segmentation - Strategies

The customer file can have several overall segmentation schemes depending on the objective to be attained. In general, they can be classified into three main categories:

- Promotional product offers
- Life-stage marketing
- Market research

# Segmentation - Strategies

## **Segmentation for Promotional Product Offerings**

The main type of segmentation is designed for promotional product offerings. Typically house customer data is used for this type of segmentation scheme. Two basic levels of segmentation are employed, each serving a unique purpose:

- Corporate level segmentation
- Product line-specific segmentation

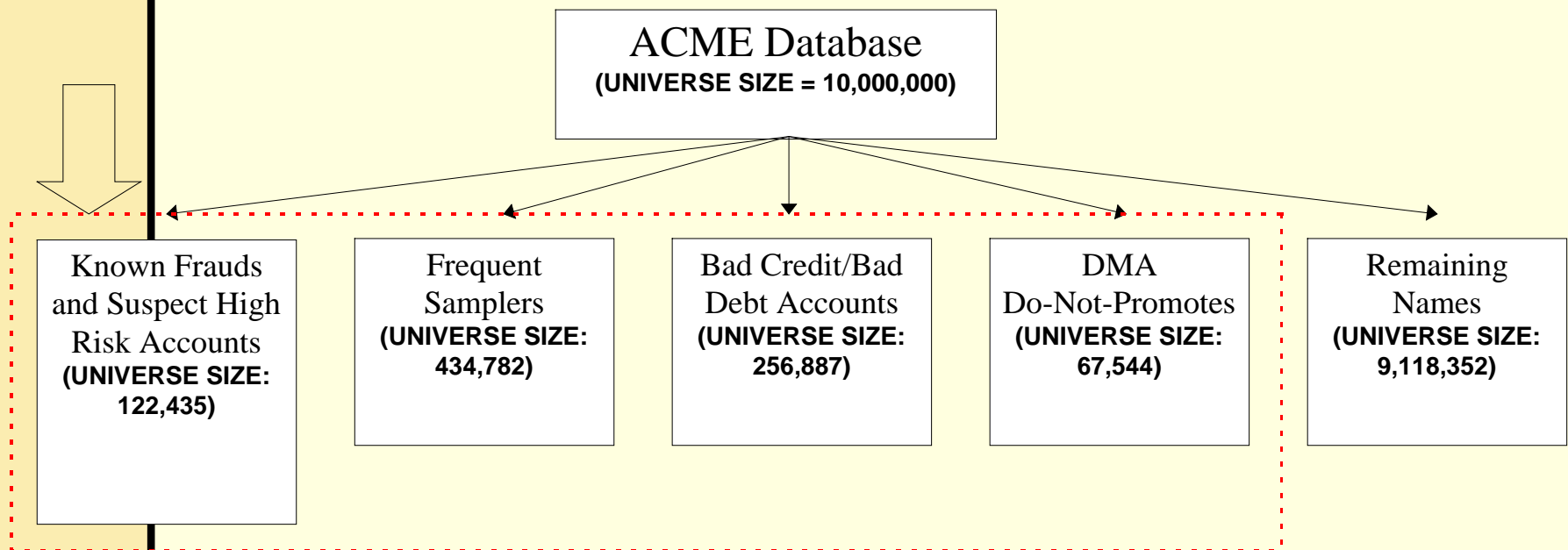
# Segmentation - Strategies

## Corporate Level Segmentation

Corporate level segmentations, sometimes referred to as corporate level eliminations, are concerned with issues that are common across all product lines within a corporation. The purpose is to inspect the entire file for names you do not want to pass along to any product line or division for promotional purposes.

# Segmentation - Strategies

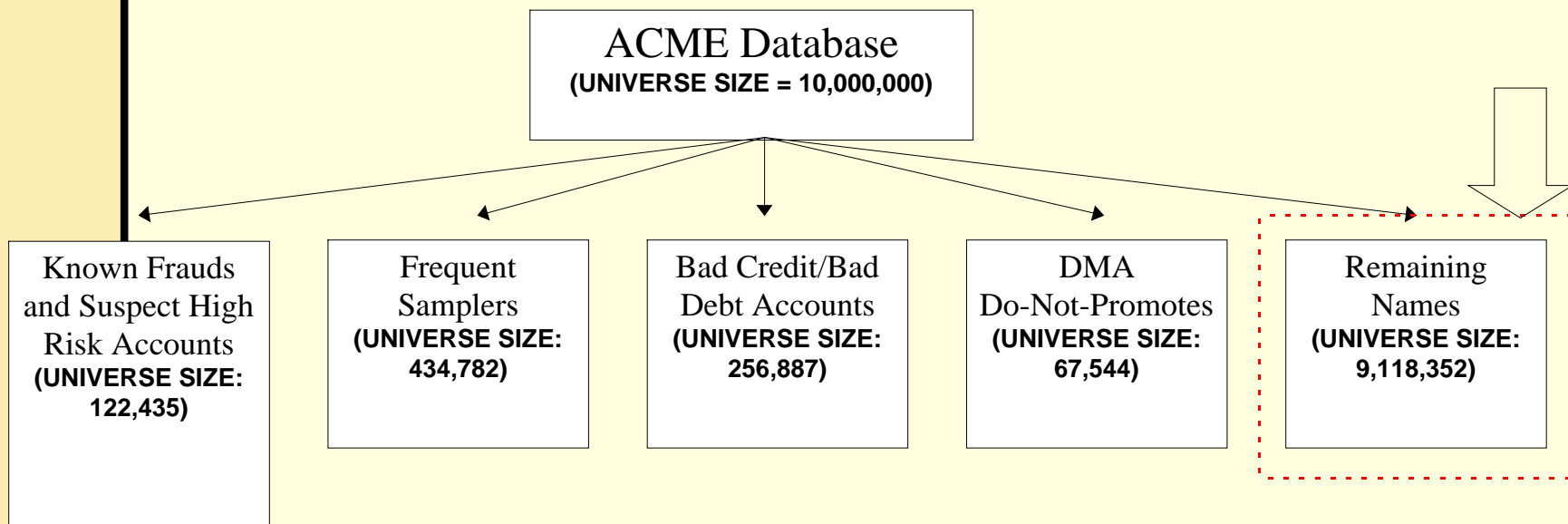
For example, common segments of names isolated at this level for direct marketing firms include DMA do-not-promotes, opt-outs, frauds, suspect high risk accounts, etc. The figure below displays how ACME Direct, a direct marketer of books, music, videos and magazines, segments their customer file of 10,000,000 customers at a corporate level.



# Segmentation - Strategies

## Product Line-Specific Segmentation

Once corporate level eliminations/segments have been completed, the next step is to determine how to segment the remaining names for each product line within the organization. Each product line will segment the “Remaining Names” to best meet their objectives.



# Segmentation - Strategies

Typically, a product line segmentation scheme divides the “Remaining Names” (typically called “promotable” names) into your most to least active groups and finally, the conversion segment. These divisions are based on recency, frequency and monetary data related to the product line of concern.

**The most active segment** is considered the most profitable and generally consists of the most recently active customers for a given product line. Product managers will *promote and test* names within this segment.

**The least active segment** is somewhat less profitable and less active than the prior segment. Product managers employ various *resuscitation* marketing tactics to ensure these names do not become purge eligible.

**The conversion segment** is defined as the group of customers who have never purchased products within the product line of interest. Product managers will employ *conversion* tactics to move these names into the core segments above.

# Segmentation - Strategies

Defining the names that belong in the conversion segment is easy. These are the people who have never purchased a product from the product line of interest.

Defining the others segments is a more complex process that requires the analysis of recency, frequency and monetary customer data to determine the correct splits

# Segmentation - Strategies

## **Segmentation for Life-Stage Marketing and Research**

Life-stage segmentation divides the file in a way that considers primarily demographic and psychographic data. This enables marketers to develop, market, or advertise more relevant products and offers on the basis of their customers' life-stages. Segmenting a customer file in this manner also allows a direct marketer to understand the future needs of their customers via research.

Data used for these types of segmentation schemes typically include individual level demographic and psychographic data as well as geo-demographic level data such as the PRIZM clusters.

# Segmentation - Strategies

Life-stages can be modeled using a multitude of internal and external lifestyle data. Life-stage segments residing on a customer database might include:

- Young families with children
- Newly moved
- Professional 25-40 year olds, no children
- Entering the retirement years
- Children 2-5, 6-8, 9-12
- Adolescents
- College students
- Empty nesters
- New grandparents

# Segmentation - Strategies

These segments can be used to determine the future needs of your customer base via research enabling you to promote appropriate and relevant promotions, offers, products or services.

A direct marketer may decide to first determine the potential size of a “empty nesters” segment by enhancing the file with data to identify such households. Next, focus groups or surveys may be conducted to these individuals if their needs are not truly understood. Depending on the results of such research, a marketing strategy may be put in place to sell such households specific products and/or services.

For example, it is a well known fact that in the financial services industry that “empty nesters” are great prospects for annuities and that “young families with children” are good candidates for insurance.

# Segmentation – Analytical Techniques

There are four commonly used analysis methods for segmenting a customer file for promotional product offers, life-stage marketing, or market research purposes:

- Univariate and cross-tabulation analysis
- Formal RFM analysis
- CHAID analysis
- Cluster and Factor Analysis

NOTE: In segmenting your customer file, you are advised to consider each of the techniques either separately or in combination to determine the best scheme for your customer file.

# Segmentation – Analytical Techniques

This morning we will discuss one method – namely CHAID Analysis.

*And, time permitting we will review a real life case study involving CHAID.*

# Segmentation - CHAID Analysis

CHAID is an acronym for Chi-Squared Automated Interaction Detection, sometimes referred to as a “tree algorithm.” Most statistical programs such as SAS, SPSS and BMDP can run CHAID analysis.

The output of a CHAID analysis will result in what is typically called a “spider chart” or “tree diagram.” All splits will be statistically meaningful in terms of maximizing separation in the data with respect to the customer action of concern (response, payment, renewal, etc.). The statistical details of how this is accomplished will not be discussed in this lecture.

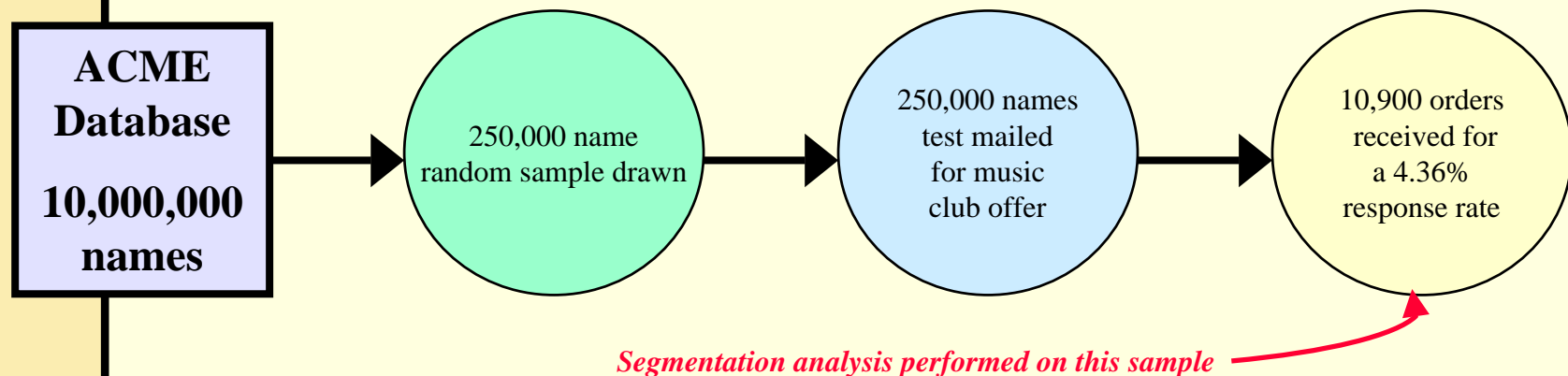
However, we will go through an example that shows how the CHAID algorithm works.

# Segmentation - CHAID Analysis

ACME Direct is a direct marketer of books, music, videos, book series, music clubs and magazines. They requires a new segmentation scheme for the music club in order to be more effective and efficient in their marketing efforts.

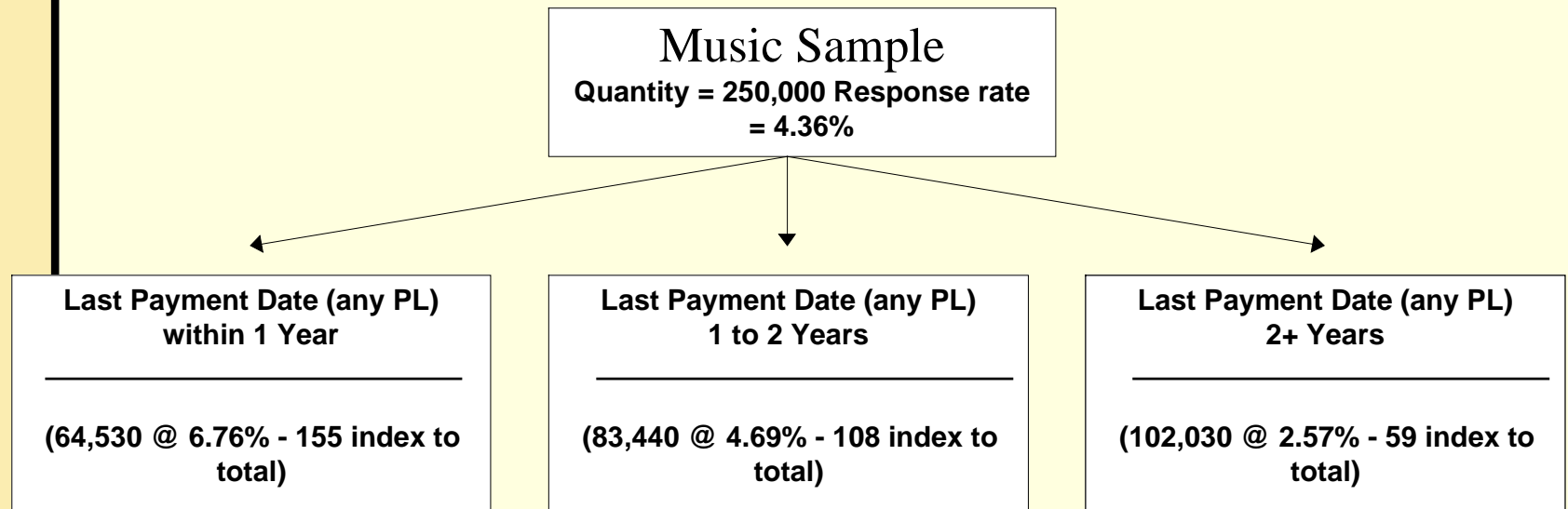
To develop this new segmentation scheme ACME decides to perform a CHAID analysis. The sample used to perform the CHAID analysis was a test mailing of one of their music club offers to the entire database of 10,000,000 records.

The total sample size was 250,000 names with a overall response rate of 4.36%.



# Segmentation - CHAID Analysis

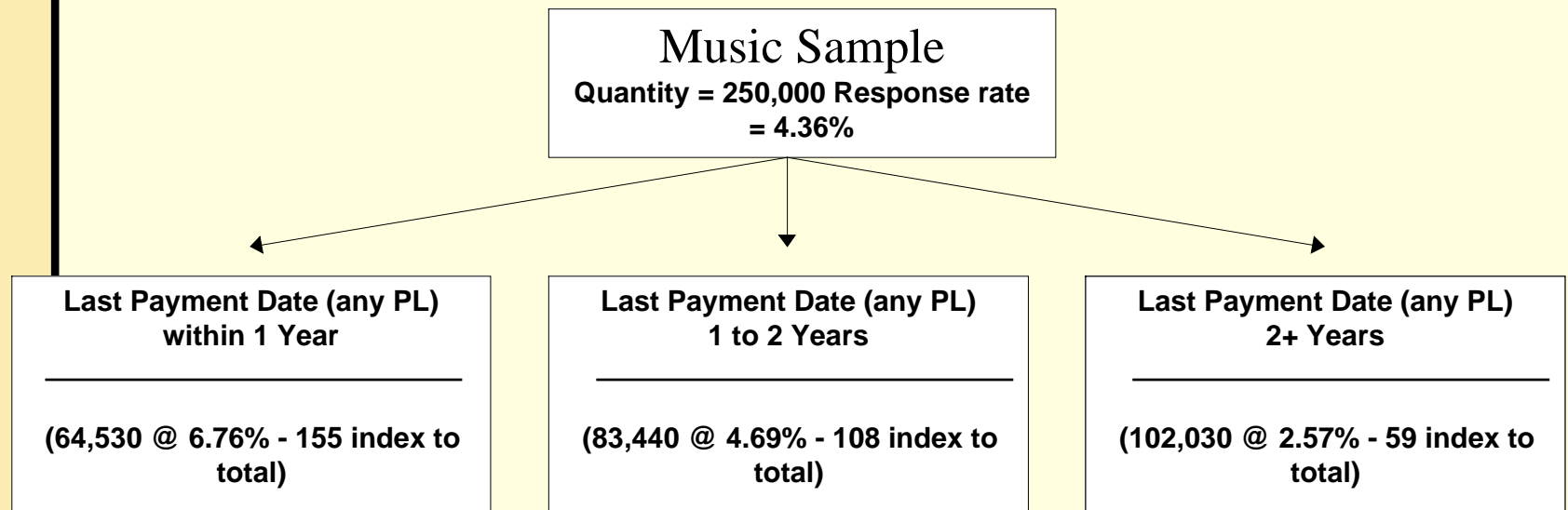
A CHAID analysis begins by examining every variable available on the sample, identifying the one variable that will most significantly maximize the separation in response rates.



In this example, CHAID determined that the variable "Last Payment Date" best discriminated between music orderers and non-orderers. In particular, three categories were generated for this variable: within 1 year, 1 to 2 years and 2+ years.

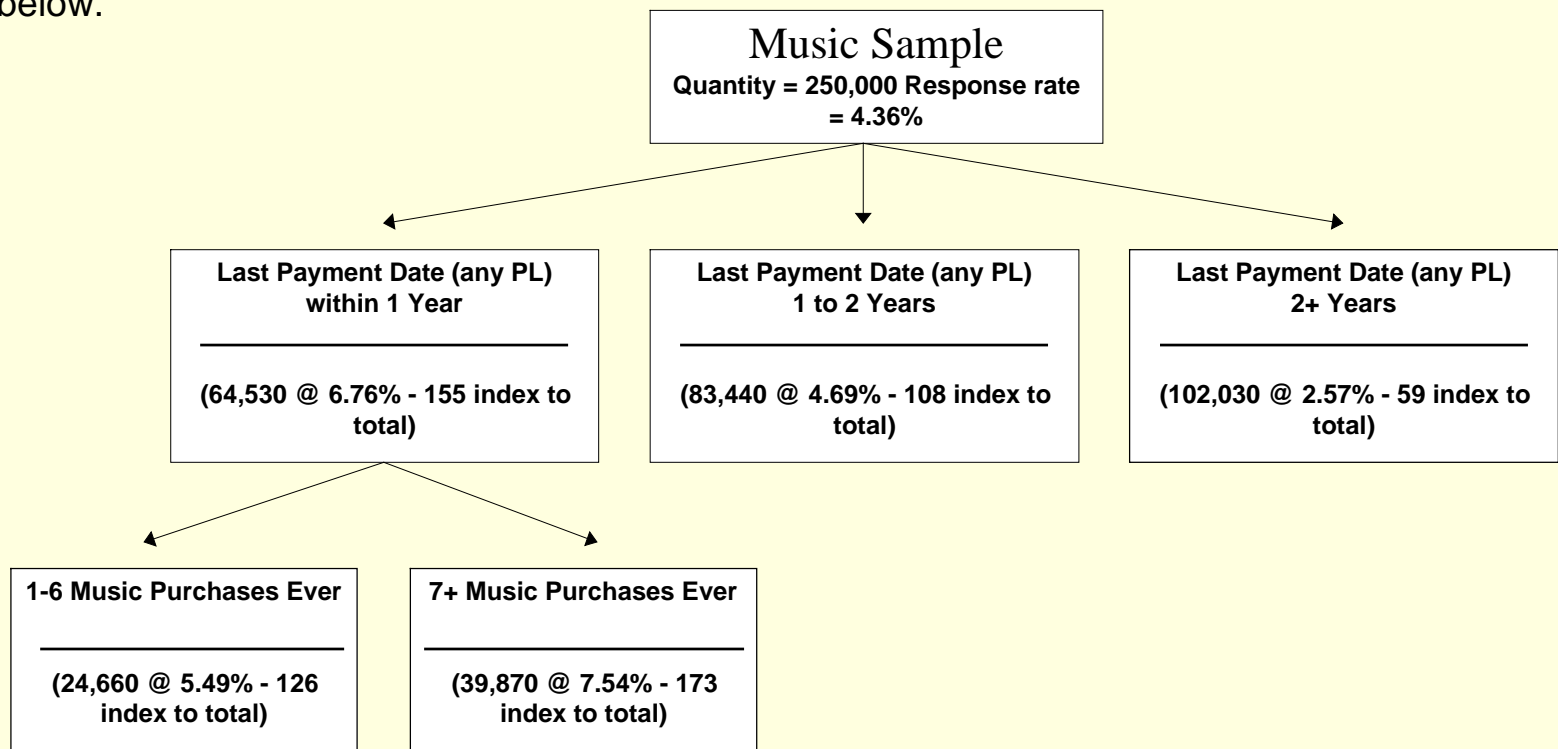
# Segmentation - CHAID Analysis

Displayed in each cell of this figure is the number of names, the response rate and the index to total. In this example, the most responsive group is those with a last payment date within 1 year. This group represents 26% (64,530/250,000) of total with an index to total of 155. The least responsive group to music promotions defined by CHAID are those with a last payment date 2 or more years ago. This group represents 41% (102,030/250,000) of total with an index to total of 59.



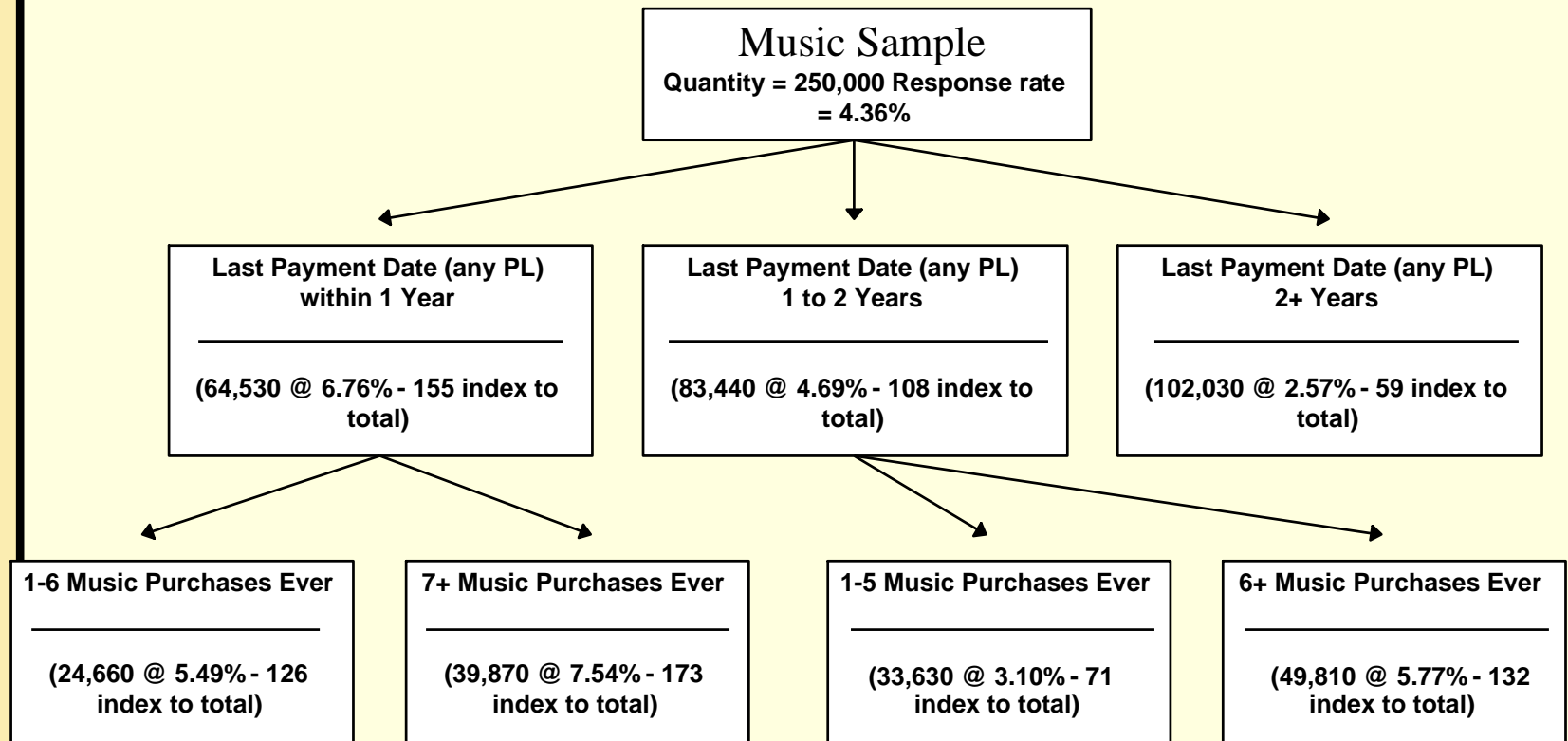
# Segmentation - CHAID Analysis

Next, CHAID examines each group at this first level individually to determine if further significant splits in the data can be made. In particular CHAID will examine the names contained in the first group “Last Payment Date within 1 Year” to determine if there is any other variable that will help further maximize the separation in response rates for these names. In this example, CHAID determined that it could further maximize separation in response for these names by using the variable “Number of Music Purchases.” Two categories were generated: 1-6 music purchases and 7+ music purchases. The resulting output is shown below.



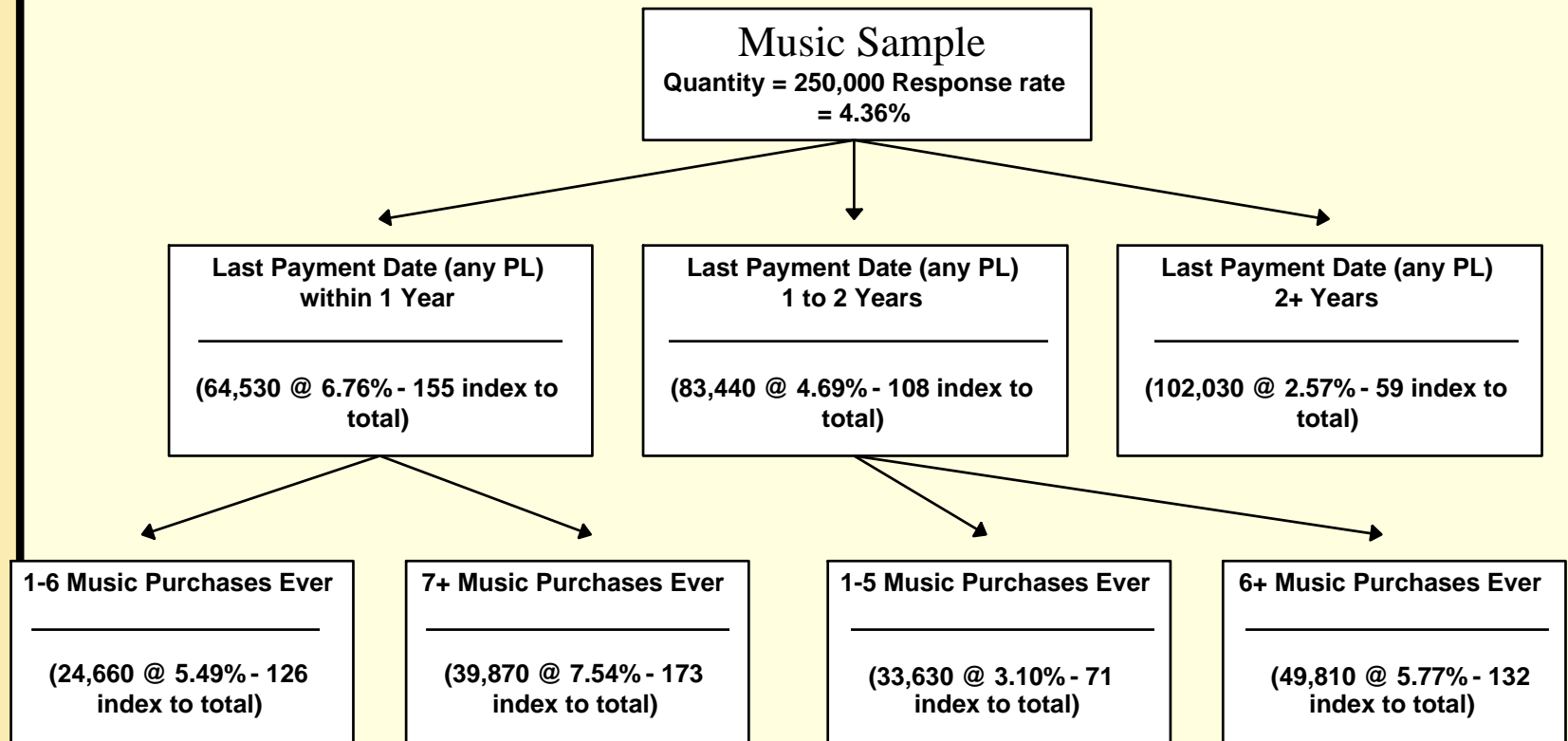
# Segmentation - CHAID Analysis

CHAID will then examine the names contained in the second group “Last Payment Date within 1 to 2 Years” to determine if there is another variable that will further maximize the separation in response rates for these names. In this example, CHAID determined that it could further maximize separation in response for these names by also using the variable “Number of Music Purchases.” However, in this case the categories created were slightly different: 1-5 music purchase and 6+ music purchases as shown in the figure below.



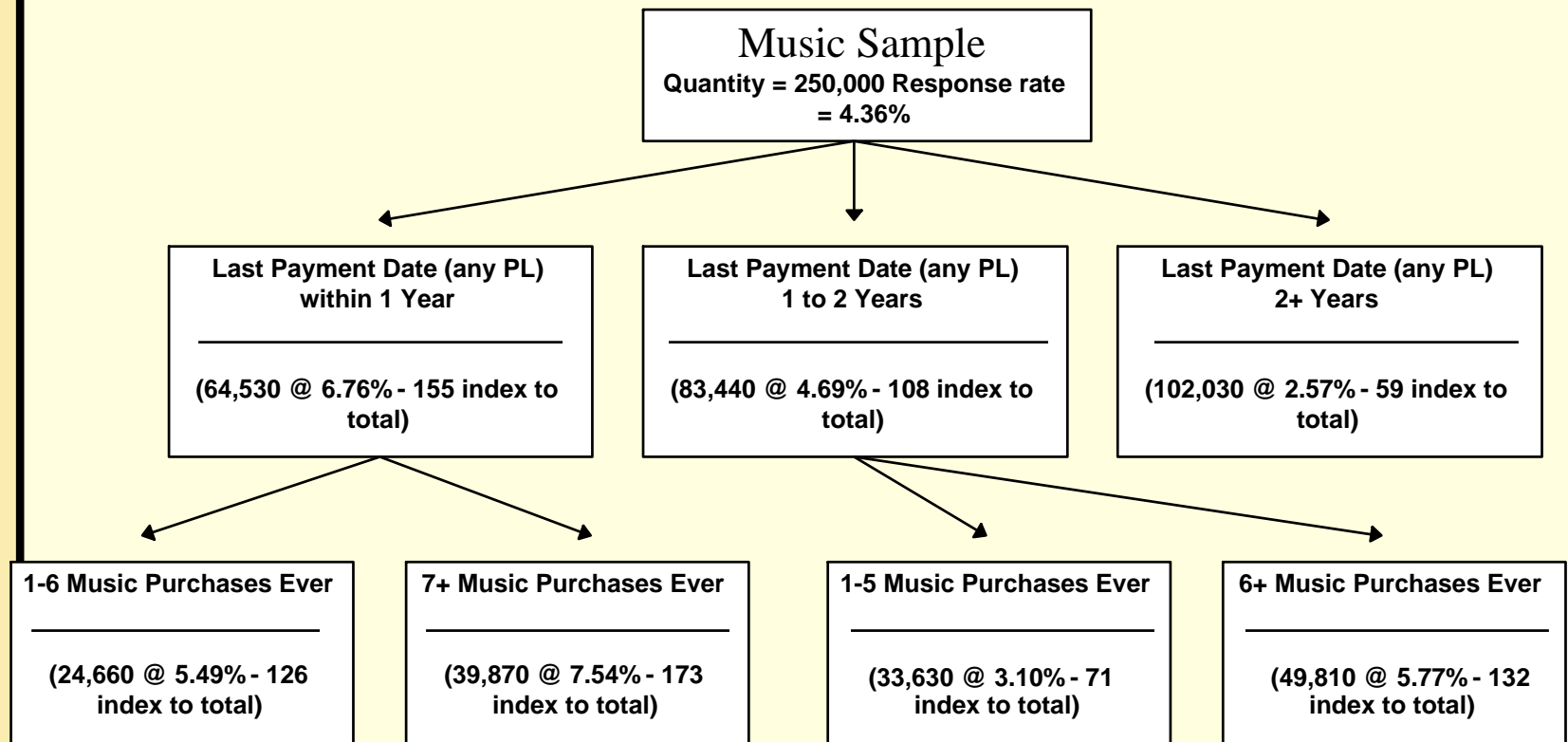
# Segmentation - CHAID Analysis

CHAID will then examine the names contained in the third group “Last Payment Date 2+ Years” to determine if there is another variable that will further maximize the separation in response rates for these names. In this example, CHAID was unable to find another variable that could further maximize separation for this group.



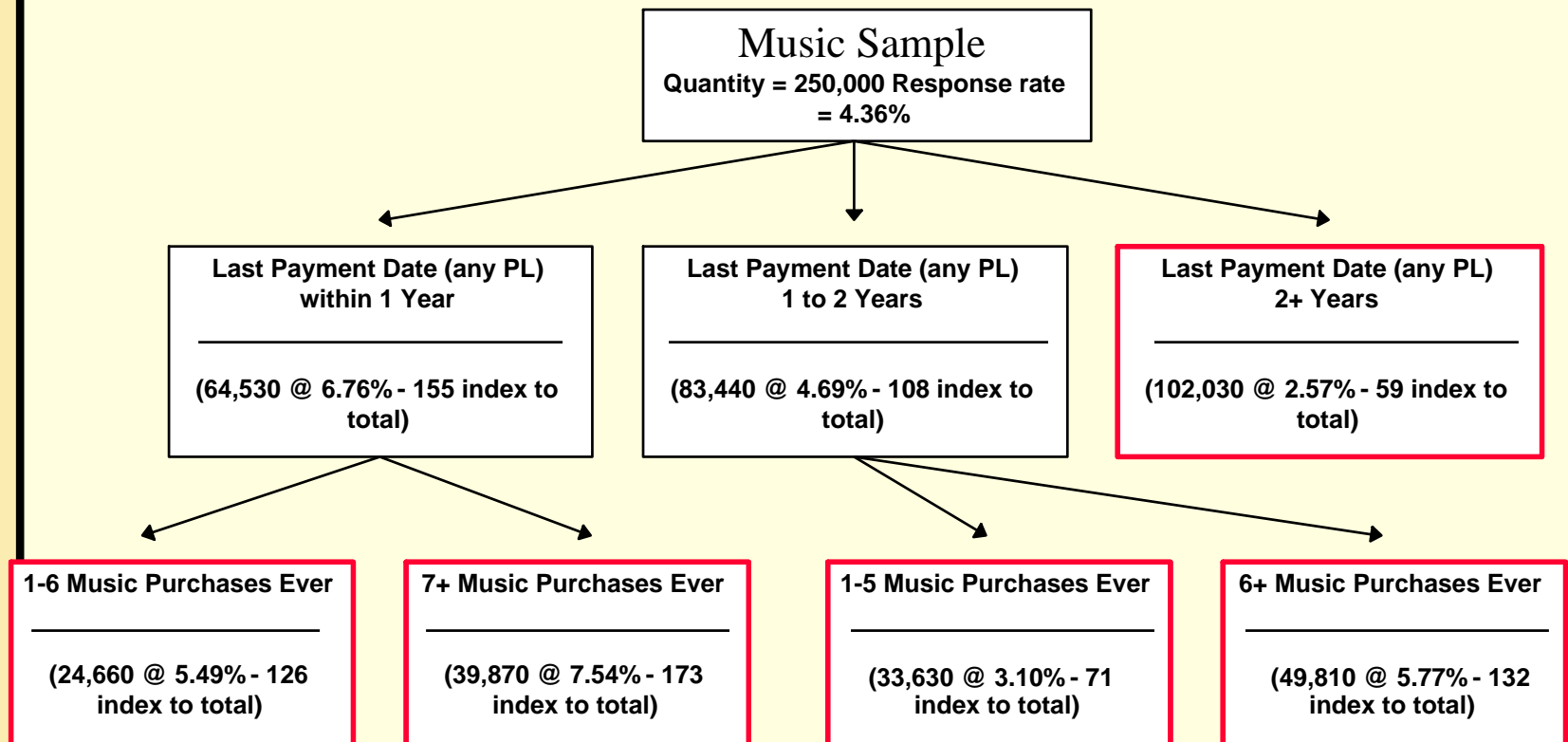
# Segmentation - CHAID Analysis

With the second level now fully defined, CHAID examines each group at this level individually to determine if further significant splits in the data can be made. In our example, no further splits were made by CHAID meaning that no other variables could be found to further the separation in response rates for any of the groups at this third level.



# Segmentation - CHAID Analysis

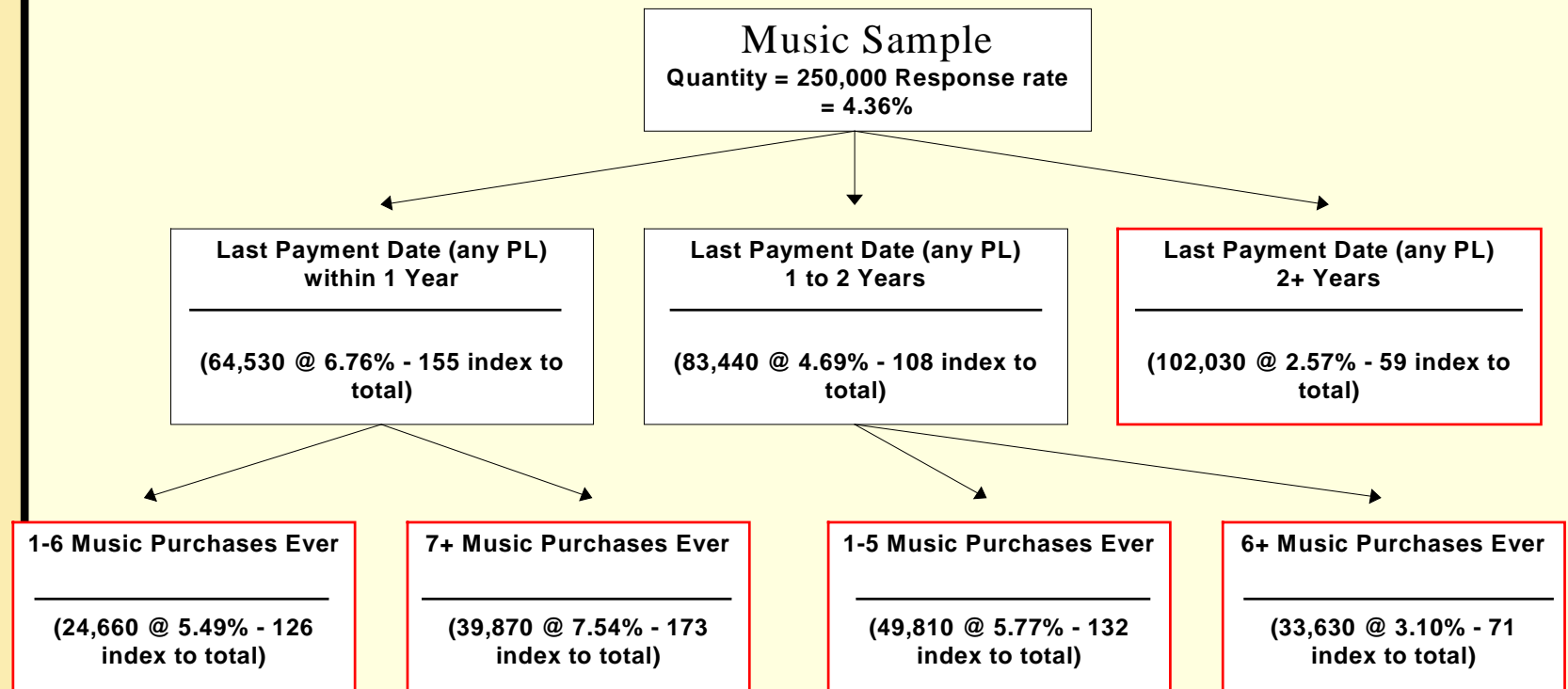
The resulting segmentation scheme produced by CHAID can be seen below. The last node or box (highlighted in red) on each branch represents the final segments for consideration.



# Segmentation - CHAID Analysis

For this example, the CHAID analysis created five segments:

- **Segment 1:** Last Payment Date within 1 year and Music Purchases Ever (1-6)
- **Segment 2:** Last Payment Date within 1 year and Music Purchases Ever (7+)
- **Segment 3:** Last Payment Date within 1 to 2 years and Music Purchases Ever (1-5)
- **Segment 4:** Last Payment Date within 1 to 2 years and Music Purchases Ever (6+)
- **Segment 5:** Last Payment Date 2+ years



# Segmentation - CHAID Analysis

The variables used by CHAID will vary depending on the business application at hand. In this case study, only “Last Payment Date” and “Music Purchases” were found to be useful for our purposes. In the telecommunications industry, for example, CHAID may consider “total services contracted” or “average monthly call volume.”

# Segmentation - Case Study Utilizing CHAID

- A seconds collection agency approached a major club continuity marketer in an attempt to try and recoup additional revenue for the marketer on unpaid accounts already worked once by another collection agency.
- Once the project was approved, the seconds collection agency took a sampling of the 1,396,449 names worked once by the other collection agency who did not pay.
- These test names were first telephone and address verified. They then received an automated voice mail message regarding their unpaid balance with the continuity marketer followed by a collection letter which included a pre-paid return envelope.
- The seconds collection agency received a 3.03% payment rate on these accounts which did meet break-even but did not meet their 10% ROI requirements.
- Since the seconds collection agency is not in the business to simply break-even, they employed the assistance of Drake Direct to determine which segments within the population were the most profitable to bill.

# Segmentation - Case Study Utilizing CHAID

Marketing cost and revenue figures for the seconds collection agency were as follows:

- Cost of automated voice message and mailer = \$1.00
- Remit rate = 10% (percent of net collected to be paid to the continuity marketer)

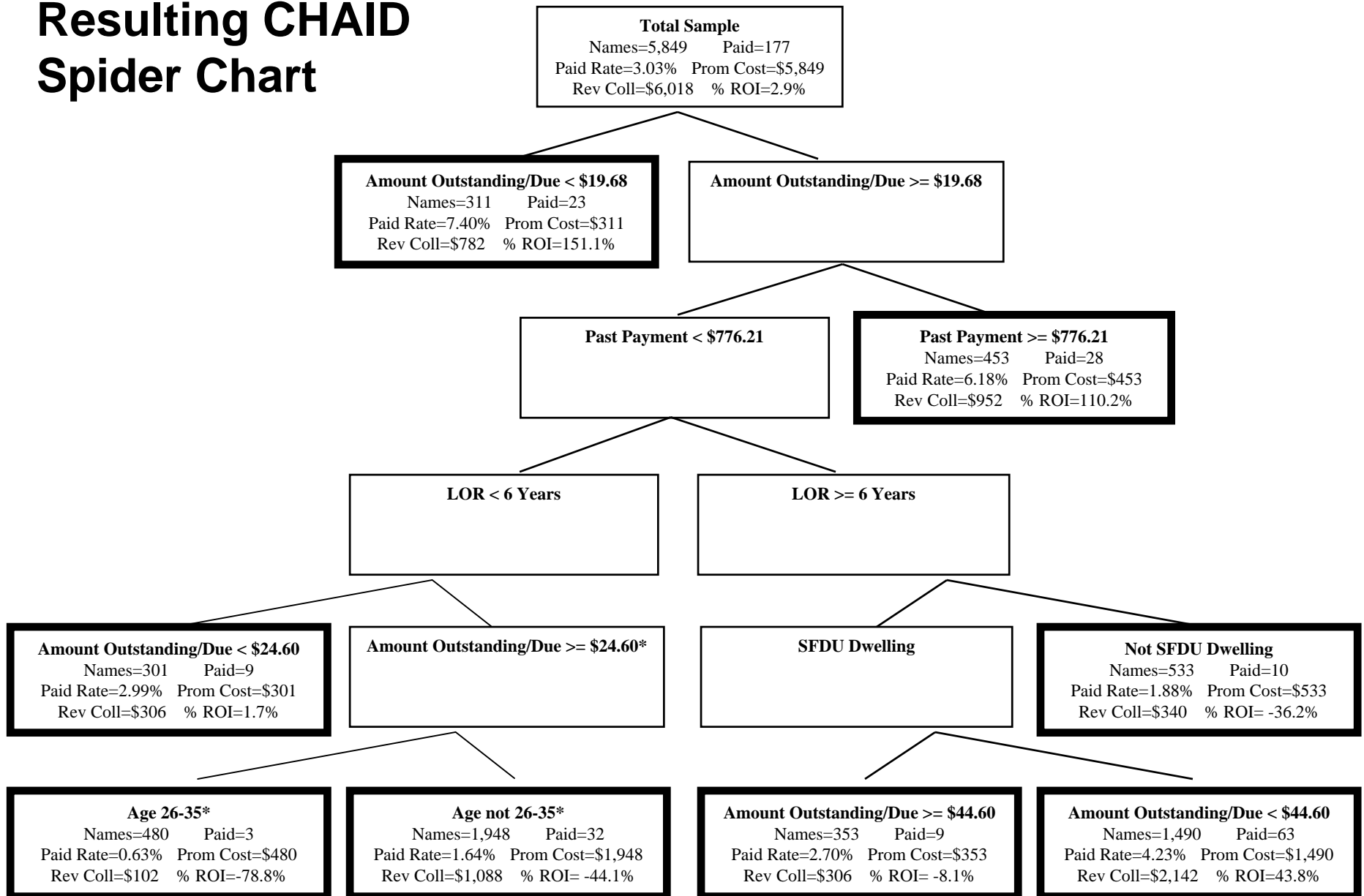
NOTE: All figures in this case study have been altered slightly for purposes of public presentation.

# Segmentation - Case Study Utilizing CHAID

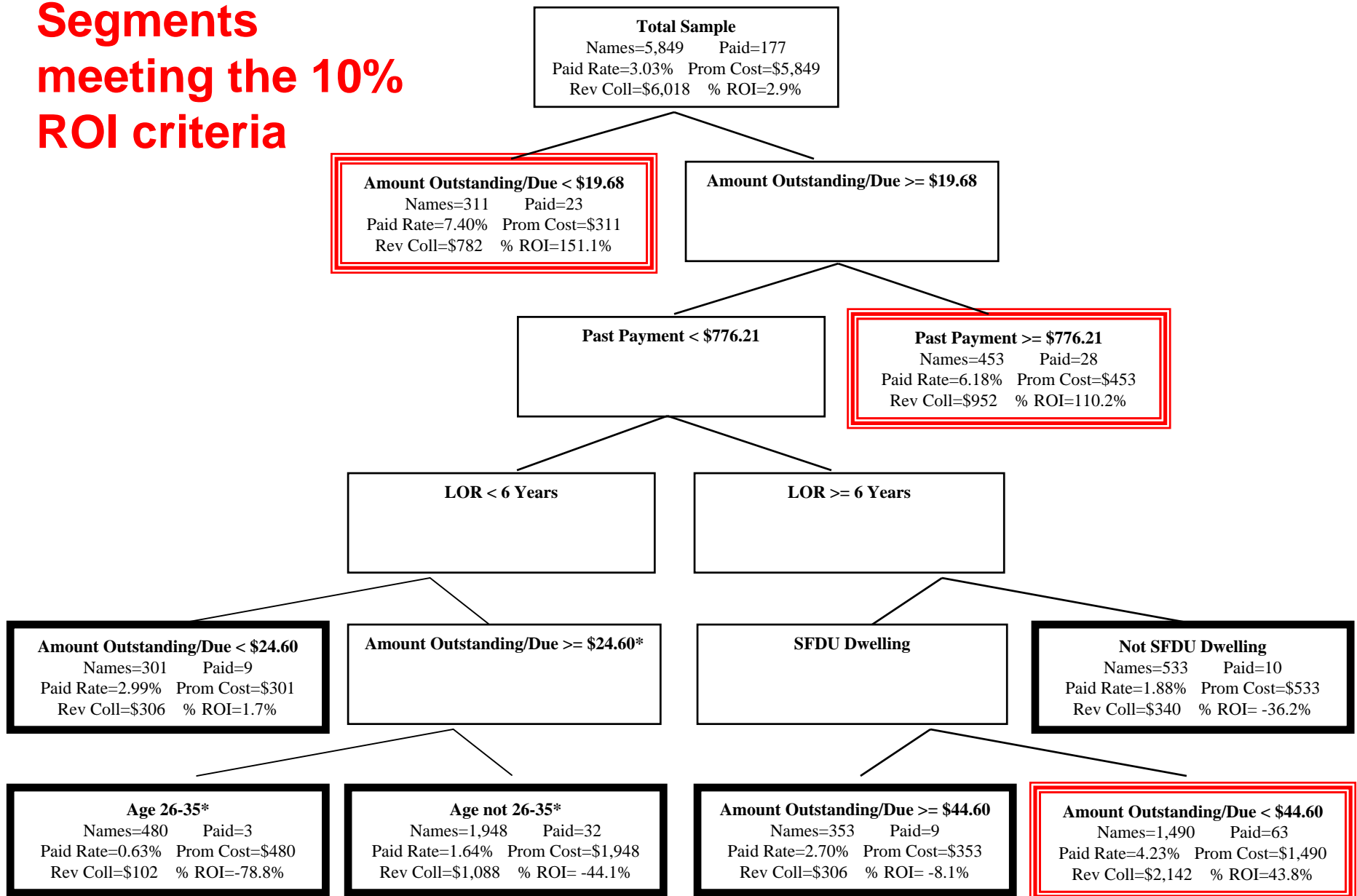
Predictor variables considered in the analysis included data from the continuity marketer such as source, affinity interests, open balance, number of days active prior going bad, total dollars paid prior going bad, etc.

In addition, enhancement data was appended to the sample including such items as age, credit card status, dwelling type, rent/own status, occupation, length of residence, etc.

# Resulting CHAID Spider Chart



# Segments meeting the 10% ROI criteria



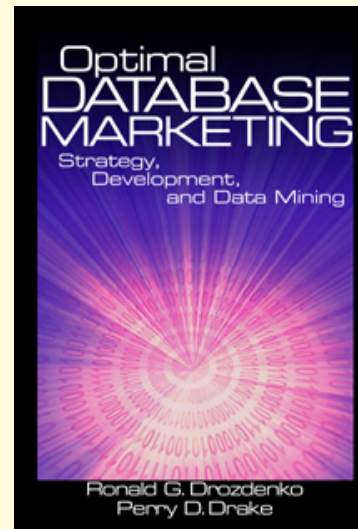
# Segmentation - Case Study Utilizing CHAID

Based on the sample findings for these three segments and when projected to all eligible names on an annual basis the following P&L results:

- Number of names eligible to be billed by the seconds annually = 1,396,449
- Estimated number of these names falling into the profitable segments = 537,633 (extrapolated from the sample results)
- Estimated number of names who will pay = 27,205 (extrapolated from the sample results)
- Cost to promote these 537,633 names = \$537,633
- Estimated revenue collected by the seconds collection agency = \$918,816 (extrapolated from the sample results)
- Net revenue prior remit =  $\$918,816 - \$537,633 = \$381,183$ .
- Remit costs = \$38,118
- Net revenue including remit = \$343,065
- Return on Investment = 62% ( $\$343,065 / \$547,633$ )
  - *well above the 10% requirement.*

## Segmentation – For More Information...

For more information on all of these topics and other techniques, see Perry's new book "Optimal Database Marketing" available at any on-line retailer including Amazon.com.



# Contact Information

## Any Questions?

If you have any question after the session, you can reach me at:

Perry D. Drake  
914-299-4879 (cell)  
Perry@DrakeDirect.com (email)  
[www.DrakeDirect.com](http://www.DrakeDirect.com) (web site)

Thank You!

**The Center for Direct/Interactive Marketing At Mercy College, Manhattan  
presents a one-day seminar on May 1st for direct marketing professionals...**



Testing Seminar Payout Chart

# **How to Improve Response and Reduce Risk By Using "Best Practices" Testing Techniques**

## **What this seminar is about...**

*In short, testing the right way.* This seminar will tell you how to set up and run your testing program and how to do it the right way.

## **Free Gift...**

**Optimal Database Marketing** by Perry D. Drake and Dr. Ron Drozdenko -- A \$54.95 value if purchased at Amazon.com

## **Your Seminar Leader...**

**Perry D. Drake**, Vice President and General Manager of Drake Direct

## **What you will learn...**

How to think about the three major test situations in direct and interactive marketing, and how the procedures vary between direct mail, print, broadcast and the internet.

### **1. Beating the control creative**

- selecting the right creative elements to test
- setting up a test design for one or multiple creative elements
- understanding acquisition versus retention strategies
- performing statistical procedures correctly

**The Center for Direct/Interactive Marketing At Mercy College, Manhattan  
presents a one-day seminar on May 1st for direct marketing professionals...**



Testing Seminar Payout Chart

# **How to Improve Response and Reduce Risk By Using "Best Practices" Testing Techniques**

## **What you will learn (Continued)...**

### **2. Finding the offer that produces better front end response or better back end performance**

- understanding the impact of the 7 elements of the offer
- setting up a test design for one or multiple offer elements
- different techniques for house versus outside databases
- results and their financial impact on the P&L

### **3. Testing new mailing lists, publications, broadcast and websites**

- how to test mailing lists and plan roll-outs
- how to test magazines and expand successful categories
- how to test TV and radio stations
- how to test the effectiveness of websites

There will be time during the day to practice what you learn by participating in class exercises using the "Plan-alyzer" software. In addition, you are encouraged to ask about special testing situations in your job.

## **When, Where and Cost...**

**Date:** Thursday, May 1st

**Time:** 8:45 - check-in and light breakfast, 9:00 to 5:30 - seminar, Noon to 1:00 - lunch on your own

**Place:** Mercy College, Manhattan, 66 West 35th Street, 7th Floor (near 6th Ave.) -- **Bring picture ID.**

**Cost:** \$495

This document was created with Win2PDF available at <http://www.daneprairie.com>.  
The unregistered version of Win2PDF is for evaluation or non-commercial use only.