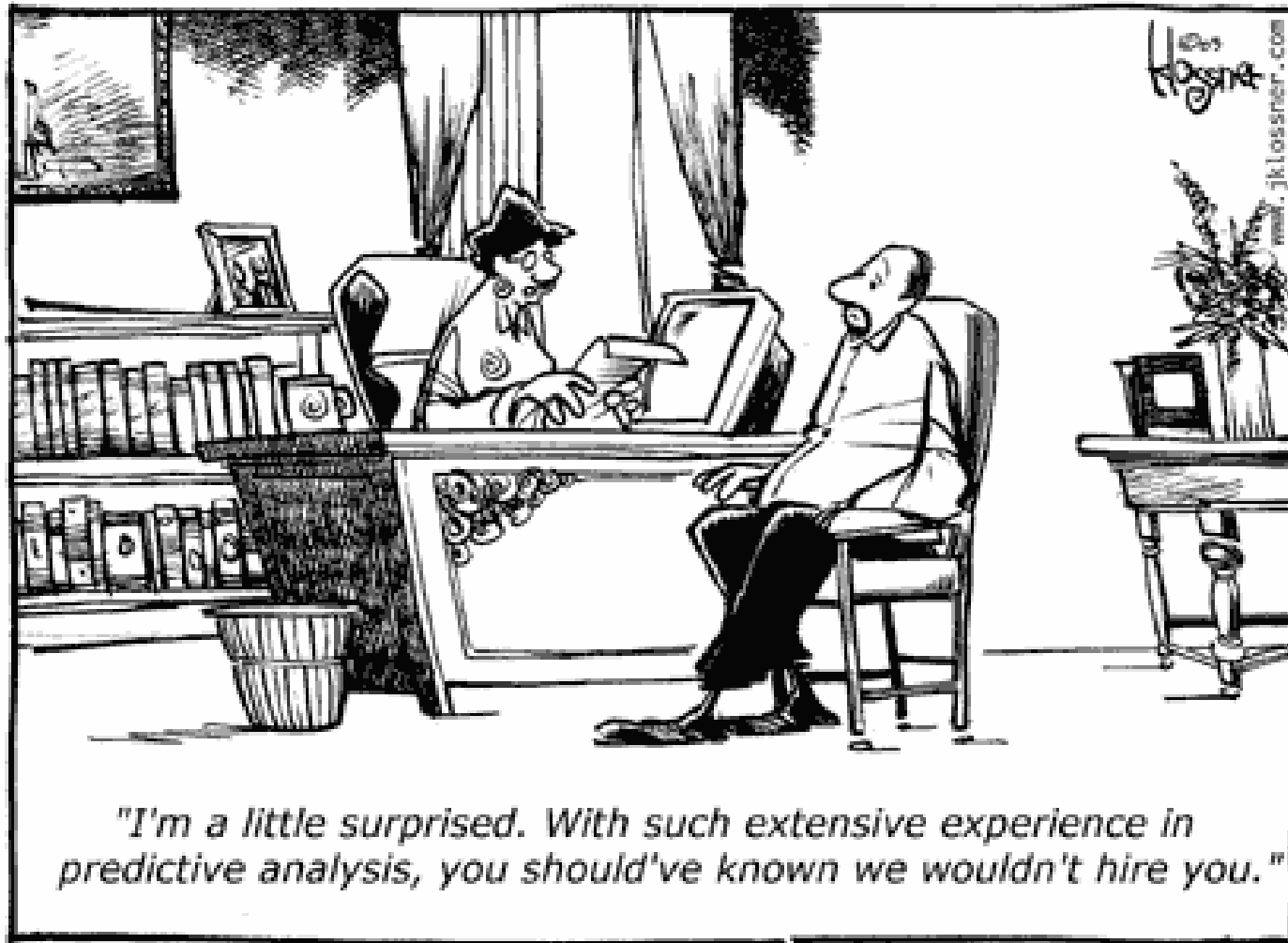# Beyond The Customer: How to Expand the Scope of Advanced Analytics

Dr Barry Leventhal
Advanced Analytics Workshop, Warsaw
25th April 2007

You've never seen your business like this before.

How far Beyond the Customer should we go?

"I'm a little surprised. With such extensive experience in predictive analysis, you should've known we wouldn't hire you."

Advanced Analytics:
Getting the Most out of your Data

Customer

Behavioral Clustering	Profitability Clustering
Attrition	Declining Spenders
Time to Attrition	Acquisition

Marketing

Store Response	Product Response
Cross Sell	Promotional Price Optimisation
Contact Optimisation	

Stores & Sales

Store Clustering	Grand Opening
Channel Modeling	Demand Forecasting

Financial

Predictive Lifetime Value	Customer Lifecycle
Credit Risk	Debt Collection

Product

|	
Individual Propensities	New Product Propensities
Product Attrition	Time to Adoption
New Product Adoption	Price Optimisation

Supply Chain

Supply Chain Forecasting

Agenda (based on your votes!)

- Pricing
  - > Price Optimization
  - > Markdown Optimization

- Product
  - > Demand Forecasting and Management
  - > Assortment and Space Optimization
  - > Product Availability

- Fraud
  - > Basic Fraud Detection
  - > Advanced Methods

- Warranty
  - > Early Warning System

Different Types of Pricing
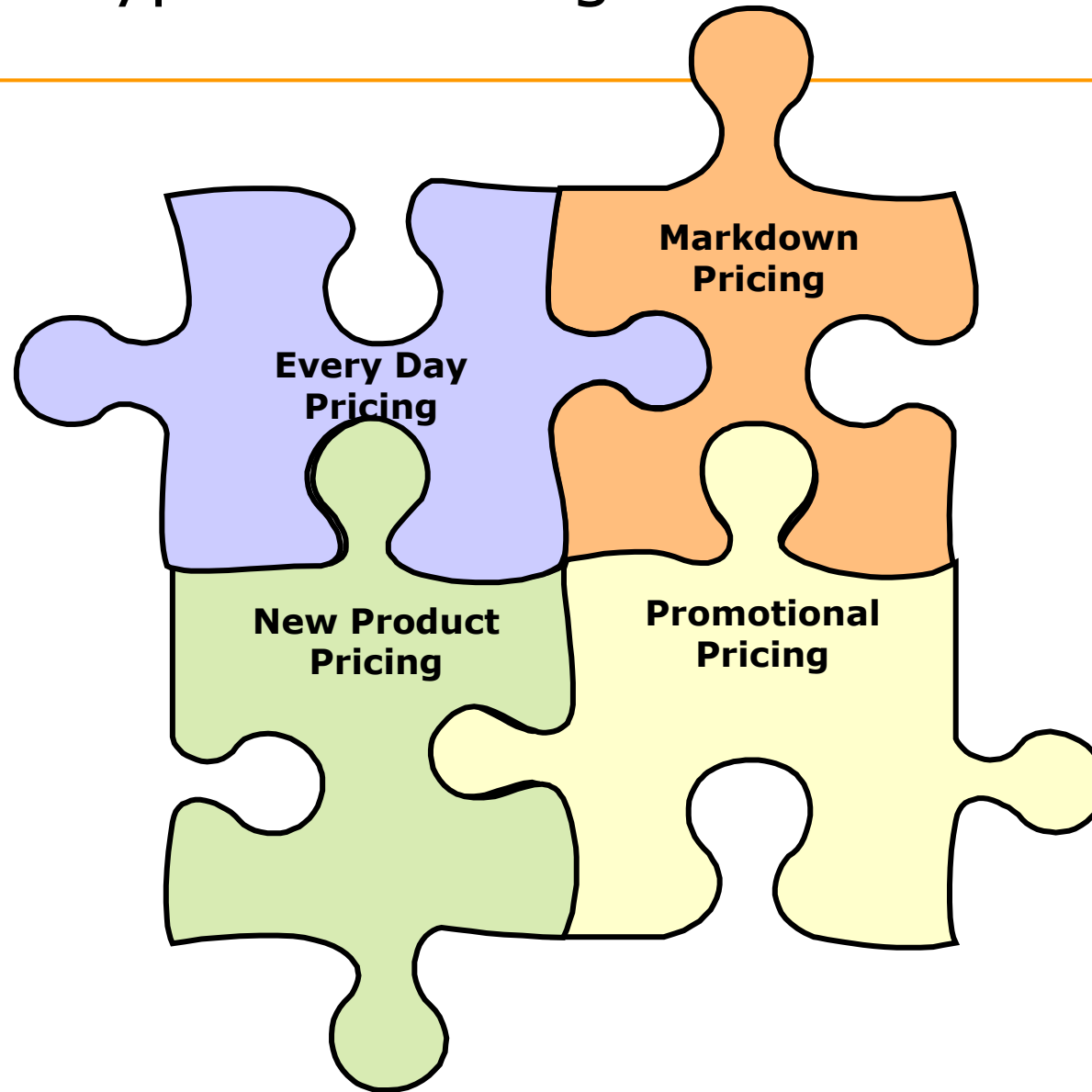
Price Optimization - The Problem

"An increasing number of shoppers say that **price trumps (wins over) brand** and the number of shoppers who say that they **stick with brands** they like continues to *decline*…

Regardless of your [retail] category, your shoppers are not that unique; **they are shopping around**. Your job is to figure out how and why they are unfaithful and build plans to address their reasons."

Forrester's "Rebuilding Consumer Loyalty" November 2003

Teradata
a division of NCR

The Opportunity

"Pricing right is the fastest and most effective way for managers to increase profits. …

1. If Price rise of 1 percent, and volumes remained stable, would generate an <u>8%</u> increase in operating profits

2. … an impact nearly <u>50%</u> greater than that of a <u>1%</u> fall in variable costs

3. … and more than three times greater than the impact of a <u>1%</u> increase in volume.

The Power of Pricing, Marn, M, Roegner, E. and C. Zawada, Mc Kinsey & Company 2003 "

Teradata
a division of NCR

BUT, the problem with deeply discounted prices

- Customers can be "trained" to expect discounted prices.
  > The Price becomes the "real" price.

- You may begin to attract the "wrong" kind of customers – those that love you just for your discounts – and they won't buy any of your non-promotional products.

- Competitors can imitate your prices.
  > That's how price wars are started.

- You may be giving away much more than you have to – even in a very competitive market.

Teradata
a division of NCR

Pricing in Retail is many different things

- **Every Day Price Optimization**
  - > Increase margin on items with low price sensitivity
  - > Increase sales of private label
  - > Limit cannibalization of strategic products and high margin products
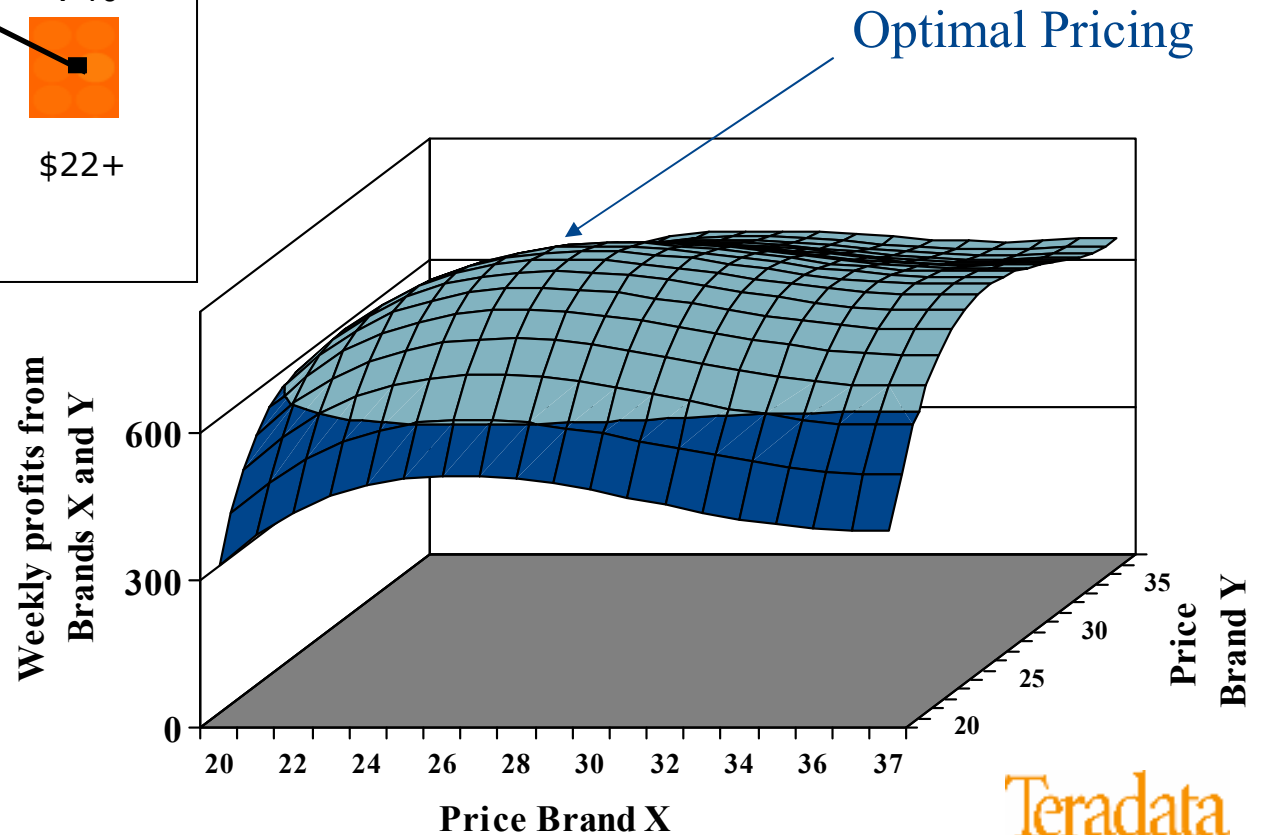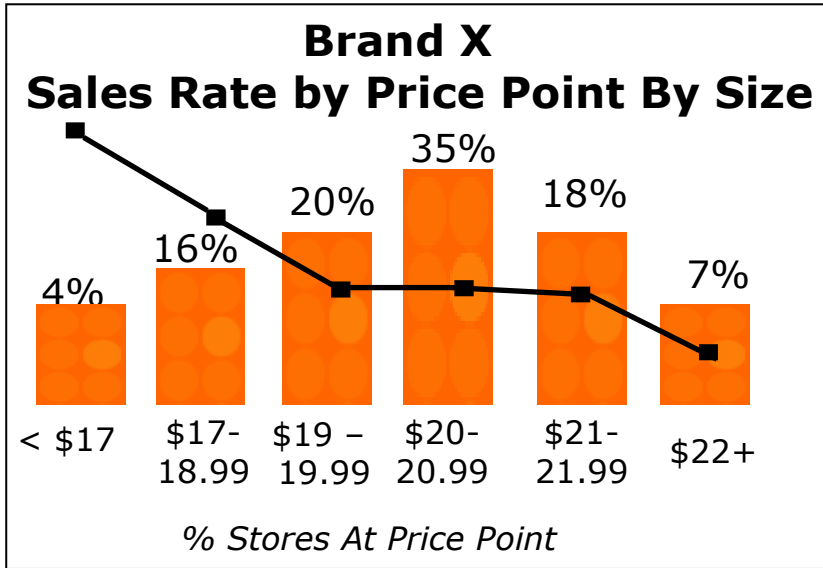 - > Remain competitive

- **Promotion Pricing**
  - > Give customers enough incentives
  - > Promote products that will drive substantial incremental sales
  - > Promote products that will increase basket sizes

- Markdown Management
  - > Better management of pricing through the product lifecycle
  - > Smoother seasonal transitions

Teradata
a division of NCR

Price Optimization in Theory

**Brand X
Sales Rate by Price Point By Size**

35%

20%

18%

16%

4%

7%

< $17 | $17-18.99 | $19 – 19.99 | $20-20.99 | $21-21.99 | $22+

% Stores At Price Point

Optimal Pricing

Weekly profits from Brands X and Y

600

300

0

20 22 24 26 28 30 32 34 36 37

Price Brand X

35
30
25
20

**Price
Brand Y**

Teradata
a division of NCR

Price Optimization in Practice

- Price optimization application - that automatically does pricing

- Supported by an initial advanced analytics project to model:
  - Price elasticities
  - Cross-sales
  - Substitution
  - Combined effects

  and determine pricing business rules

- Automatically finds the best prices for products

- Appropriate for retailers and catalogers – through any sales channel

- Application supports regular and promotional pricing

How to Make More Money (1)

Business Need	Solution	Business Benefit
Meet market demand FAST!	Make new models on demand	React to competitor actions overnight
Discover new X-sell potential	Availability-based affinity analysis	Quantify X-sell potential
Discover new cannibalization patterns	Availability-based affinity analysis	Quantify impact of cannibalization on profit
Improve overall margin	Combine price elasticity, cannibalization, and X-sell effects into price recommendations	Complete models allow improved overall profits

Teradata
a division of NCR

How to Make More Money (2)

Business Need	Solution	Business Benefit
Adapt prices locally	Make models by region	Improve margin for each region
Adapt pricing by customer segment	Make models by customer segment	Promos by segment to boost profits
Preserve pricing knowledge	All models are transparent and remain in the warehouse	Leverage learnings over time to new pricing situations

Teradata
a division of NCR

Pricing in the Telco Sector: Tariff Optimization

- Price new tariffs
  - > Measure price parity vs. existing plans
  - > Fine-tune pricing for each market segment

- Benchmark tariffs vs. competitors
  - > Ensure new tariffs are competitive
  - > Measure price position by segment
  - > Measure impact of new competitor tariffs

- Measure impact of new plans on current revenue

Teradata
a division of NCR

Tariff Check-up

- Simulation that proposes ideal tariff plan for each customer

- Based on re-rating all call records against all plans

- Used to support sale of appropriate tariff plan to each customer

- Benefits are advertised to customers

Teradata
a division of NCR

Bundled Pricing of Mobile Services (1)

- "Bundled pricing is a powerful technique for driving mobile voice usage and kick-starting much-needed new revenue from mobile data services"

- "Affordable bundled pricing has been fundamental to mobile operators in the USA achieving average voice usage that is over five times higher than in Western Europe"

- "Bundles can encourage users to adopt and use more data services than they would if the services were purchased individually"

- "By establishing ongoing communication with customers and offering incremental refinements to their service bundles, mobile operators can strengthen loyalty and may be able to migrate customers to higher-value bundles over time"

Source: 'Pricing Mobile Services for Success: towards a bundled future'
Dr Alastair Brydon and Dr Mark Heath.

Teradata
a division of NCR

Bundled Pricing of Mobile Services (2)
7 Golden Rules

- consider the individual characteristics of each service to achieve effective pricing

- review the opportunities to make bundled pricing a core feature of service pricing

- confine flat-rate pricing to services that will not incur excessive costs for them, or use strategies to constrain heavy usage

- develop bundles of services to suit different levels of affordability

- develop bundled pricing that appeals to specific target market segments

- continually review and modify bundled pricing, as part of ongoing customer relationship management

- implement highly effective billing, service management and customer relationship management systems

Source: 'Pricing Mobile Services for Success: towards a bundled future'
Dr Alastair Brydon and Dr Mark Heath.

Teradata
a division of NCR

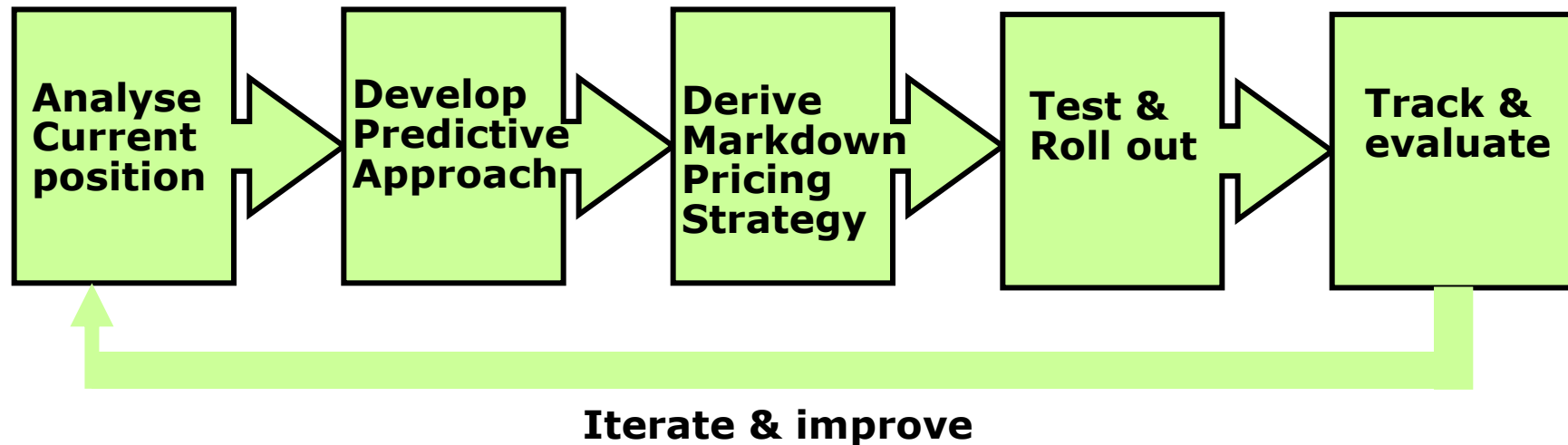# The need for Markdown Optimization in Retail

- The bottom line for retailers is critically impacted by the stock initially ordered and subsequent markdown pricing decisions

- Decisions may be challenging due to uncertainty over demand – leading buyers to overbuy and use clearance markdown to mitigate losses

- Short selling periods compound the problem, particularly if merchandise cannot be sold at the end

'Markdown' is a well known business problem

- Maximising revenues from products with limited 'shelf life'

- And minimizing losses arising from price reductions and wastage

- Significant issue in many sectors, e.g.
  - > Retail – e.g. perishable foods, fashion
  - > Hotel rooms
  - > Airline tickets

Overall steps for Markdown Optimization

Analyse Current position → **Develop Predictive Approach** → **Derive Markdown Pricing Strategy** → **Test & Roll out** → **Track & evaluate**

Iterate & improve

A Markdown Optimization application is required for most effective implementation

Teradata
a division of NCR

Deriving the Markdown Pricing Strategy

Teradata has a well defined analytical process. The steps of a markdown study are outlined below.

Markdown Analysis Steps

- Cluster stores into groups

- Identify baseline sales rates for the relevant products in each cluster

- Compute an expected sell thru curve for the upcoming period based on prior history. If a new product, use a reference product or category

- Compute the price elasticity of the products. If a new product, use a reference product or category

- If an existing product, compare sales rates against the date by which stock must be fully sold. Adjust price using elasticity.

- Monitor regularly and adjust price further as needed to put sales on pace to sell out by target date

60 – 80% of project time spent in these phases

Business Understanding

Data Understanding

Data Preparation

Deployment

Data

Modeling

Evaluation
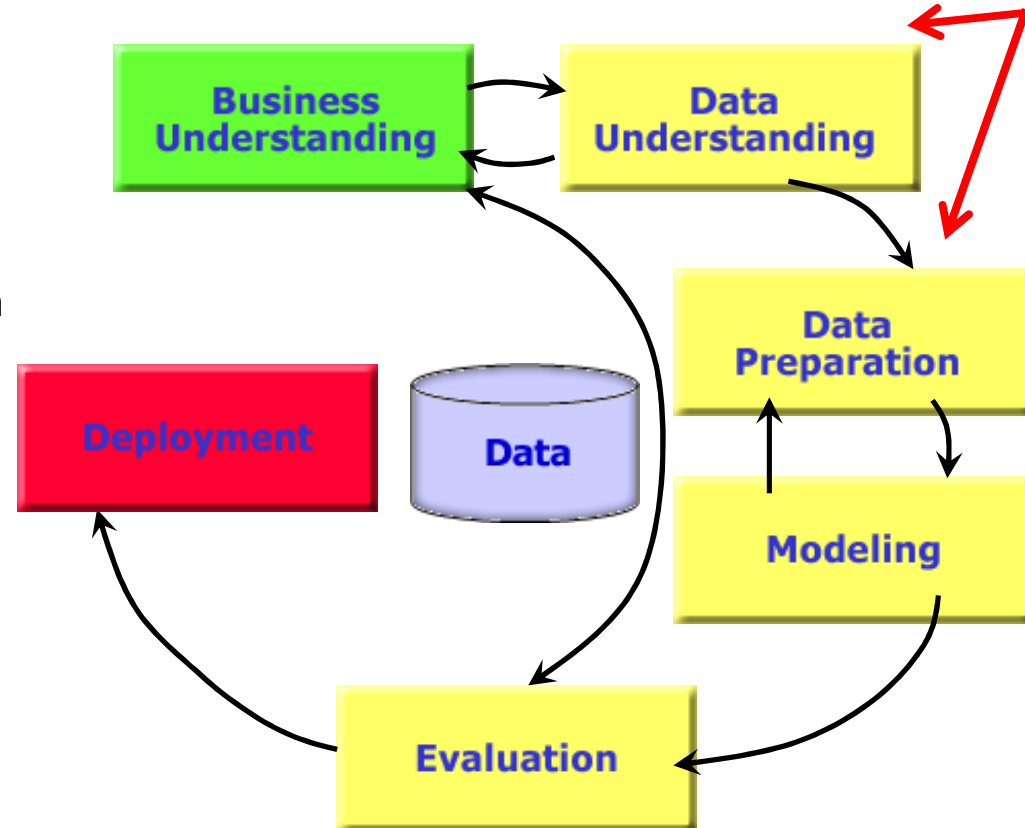
CRISP-DM data mining industry standard

Teradata
a division of NCR

Pricing Analytics - Discussion

- Have you applied these techniques?

- What are your views and comments?

Product Analytics: Overview of the Advanced Retail Planning Process

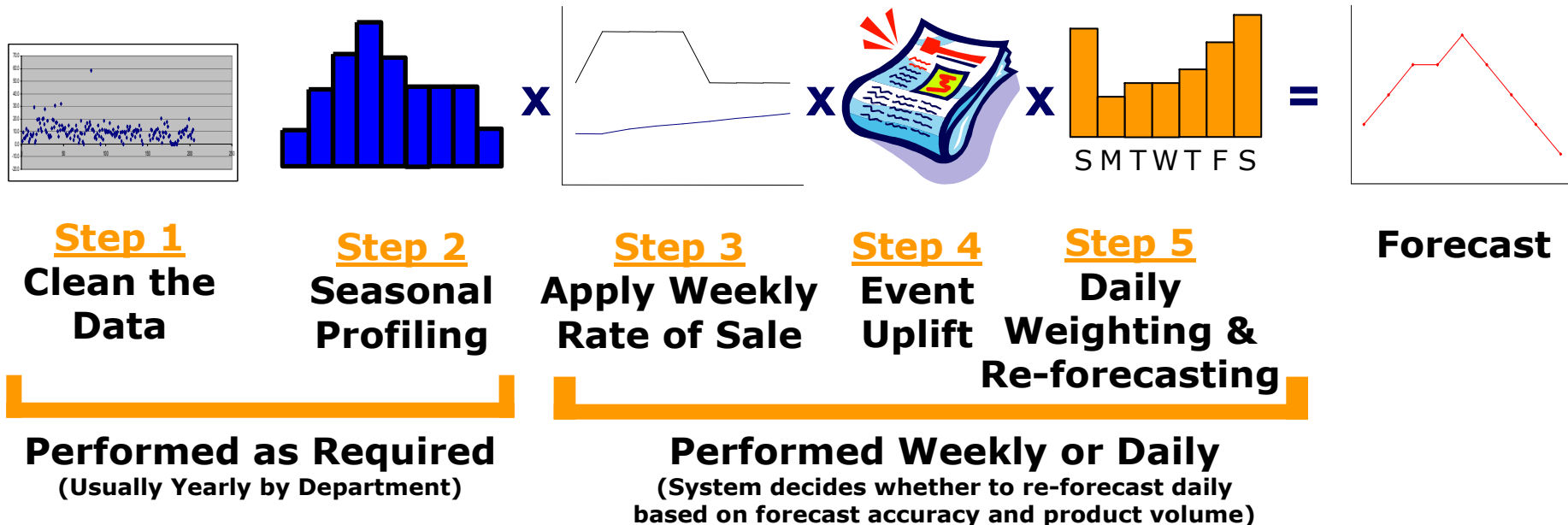**Figure 2:** Advanced retail planning process overview
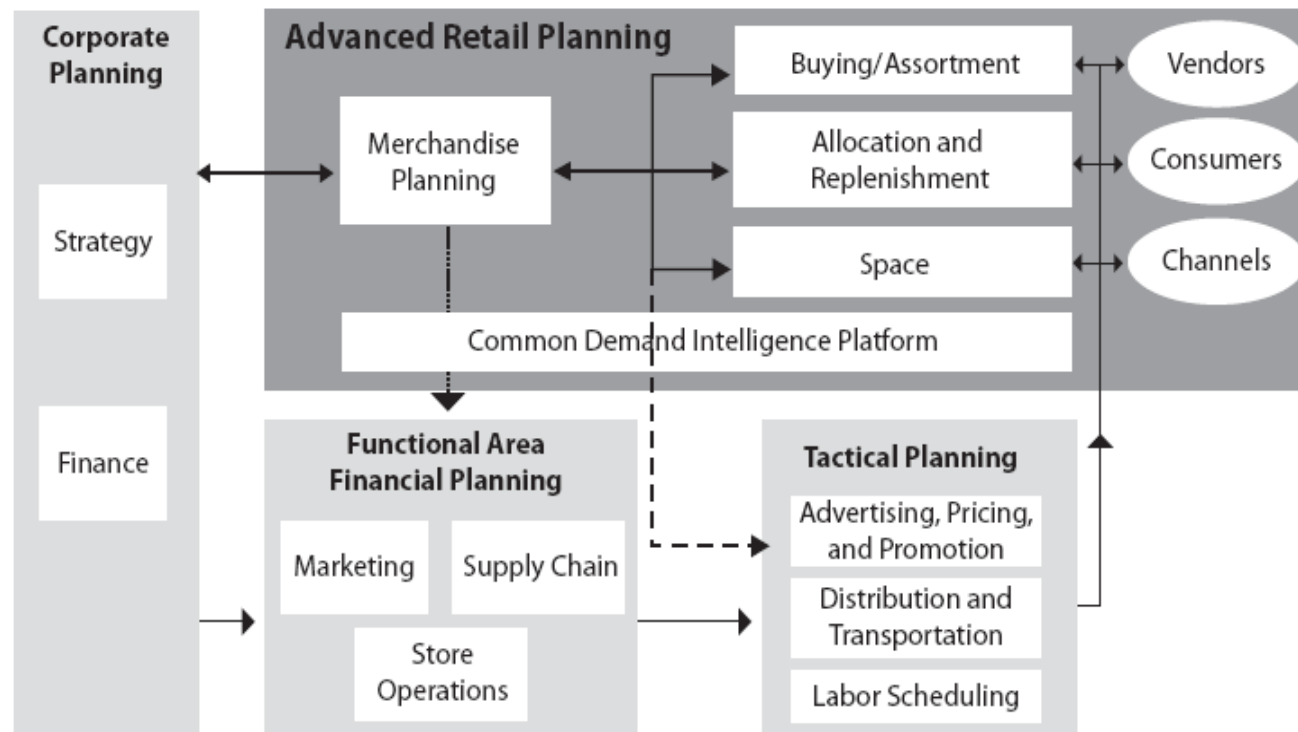
Industry analysts advocate a bottom-up approach to forecasting and planning

- AMR Research January 2005 - Striving for Advanced Retail Planning (ARP)
  - > "Traditional top-down planning does not synchronize resources for demand, due to inadequate process synchronization, insufficient analytical rigor (forecasts), cumbersome workflows"

  - > "ARP is blending detailed demand intelligence (forecasting) into synchronized assortment, inventory and space planning execution"

  - > "Advanced retail planning includes three key enhancements"
    - Science based planning using detailed demand and forecast data – bottom up perspective
    - Pervasive demand intelligence – automated demand forecasts influenced by promotions, etc..
    - Integrated sales and operational planning process driven by single unified demand forecast

Teradata
a division of NCR

Demand Forecasting Methodology
- at Store-SKU level

- Framework for Time Series Forecasting

| **Step 1**
Clean the Data | **Step 2**
Seasonal Profiling | **Step 3**
Apply Weekly Rate of Sale | **Step 4**
Event Uplift | **Step 5**
Daily Weighting & Re-forecasting | **Forecast** |

Performed as Required
(Usually Yearly by Department)

Performed Weekly or Daily
(System decides whether to re-forecast daily based on forecast accuracy and product volume)

- Other Forecasting Considerations
  > Slow-moving products
    - Generalized Autoregressive Conditional Heteroskedacicity (GARCH)
      (Reduces forecast volatility)
  > Synchronizing Store-DC-Vendor Demand
    - Store sales are translated to orders through business rules (Called Policies in DCM) like Minimums/Maximums, Lead Time, Review Time, Safety Stock, etc.
    - DC demand is aggregated and translated to vendor orders through business rules (Policies) like Minimums/Maximums, Lead Time, Review Time, Safety Stock, etc.

Advanced Retail Planning:
Store-SKU Forecast Accuracy

- Forecast Accuracy Delivers:

- Inventory productivity improvements by automatically shifting inventory investment from "C" items that are not contributing to the business to "A" items that are contributing

- Reduced markdowns through better pre season and in season buying quantities

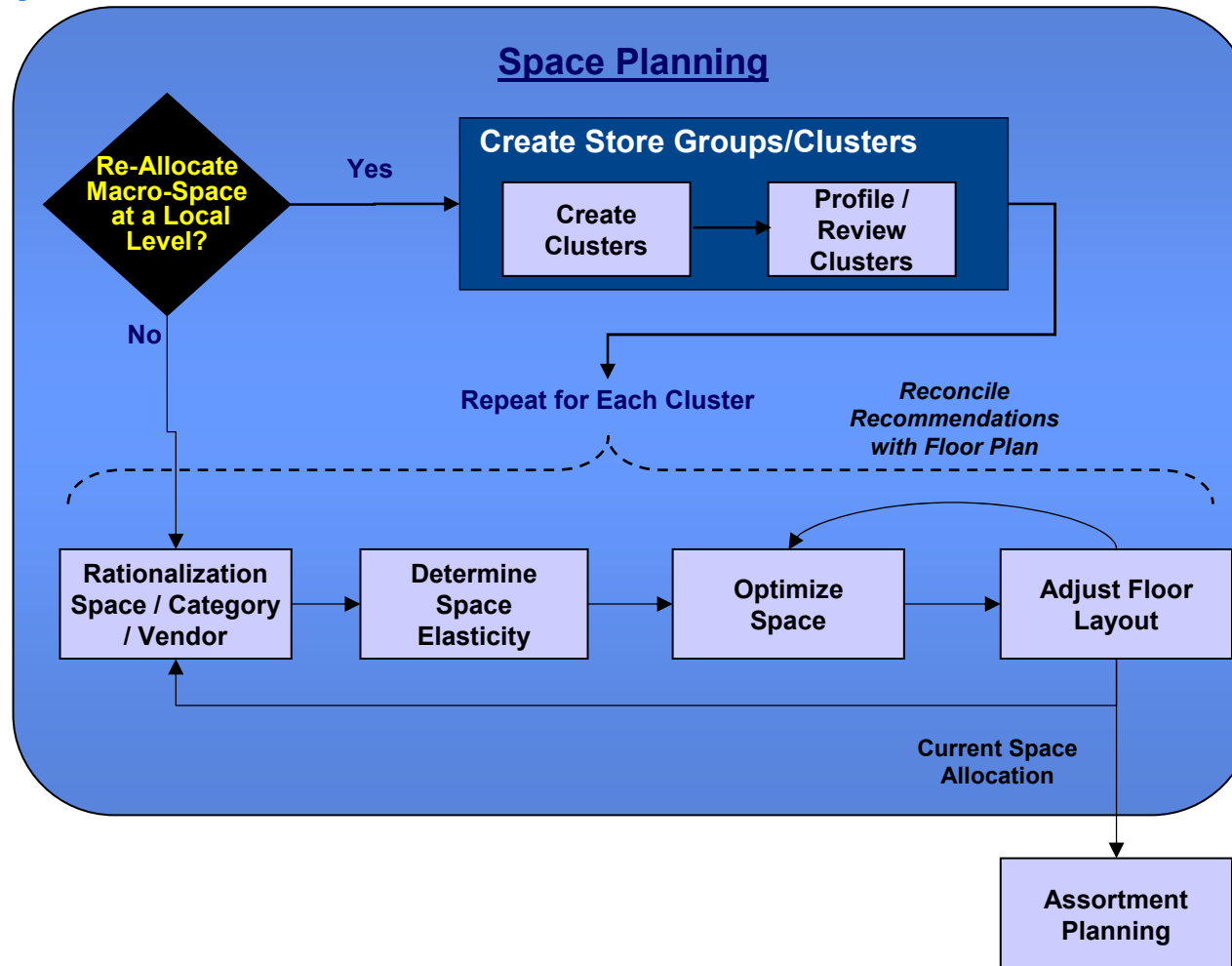- Improved customer service level and reduction of lost sales

Assortment and Space Optimization (ASO)

ASO determines both high-level space strategies and specific item presentation by store
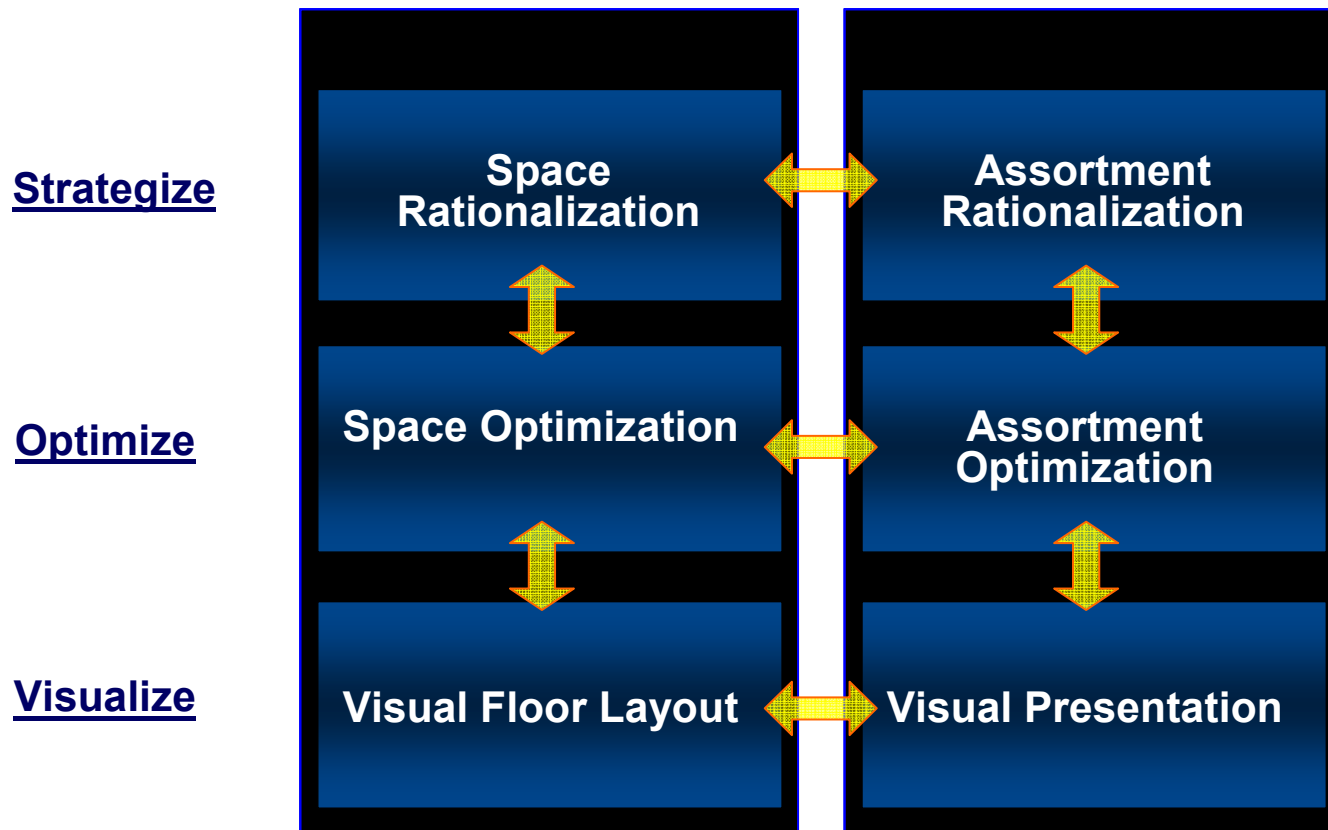
Space Planning

Space Planning uses rigorous, fact-based analysis to determine optimal category space allocation at the individual store level

Space Planning - Space Elasticity

Space elasticity is determined using regression analysis to understand how categories perform at different space allocations and/or planogram lengths

Category Profitability by Planogram

Illustrative
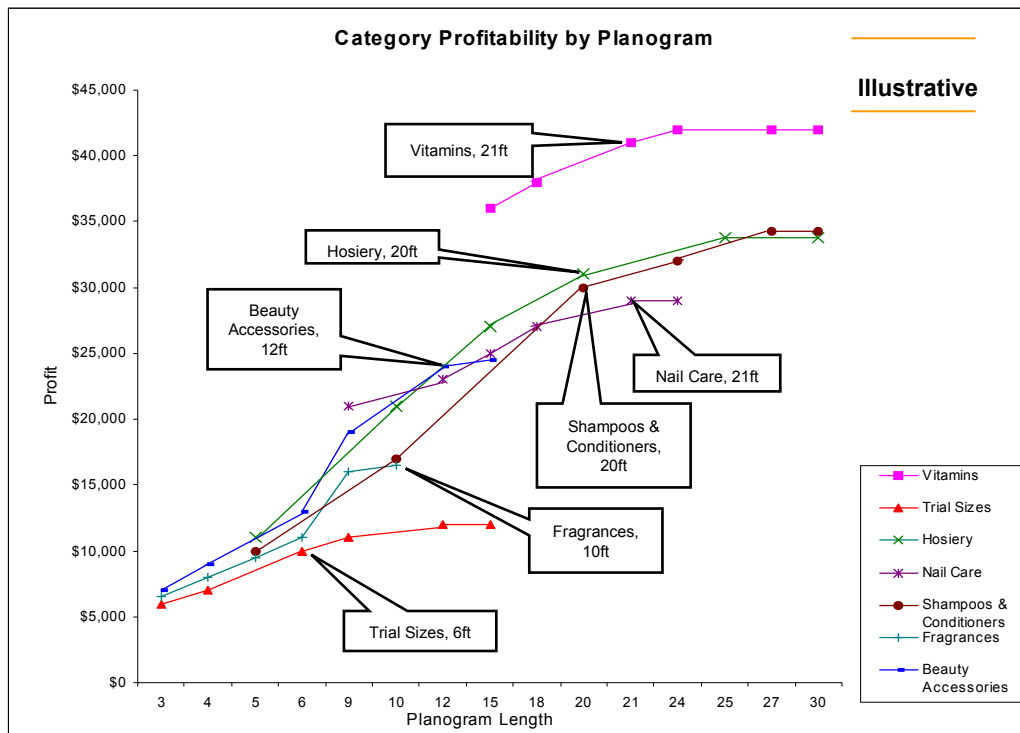
- **The analysis gives the retailer insight to incremental returns on category space investments**
- **The analysis can be performed on a chain-wide basis or within store peer groups**

Assortment Planning

Assortment Planning helps the retailer determine the ideal assortment mix to maximize sales/profit by determining which products to carry and optimizing the amount to be merchandised in each store/store cluster
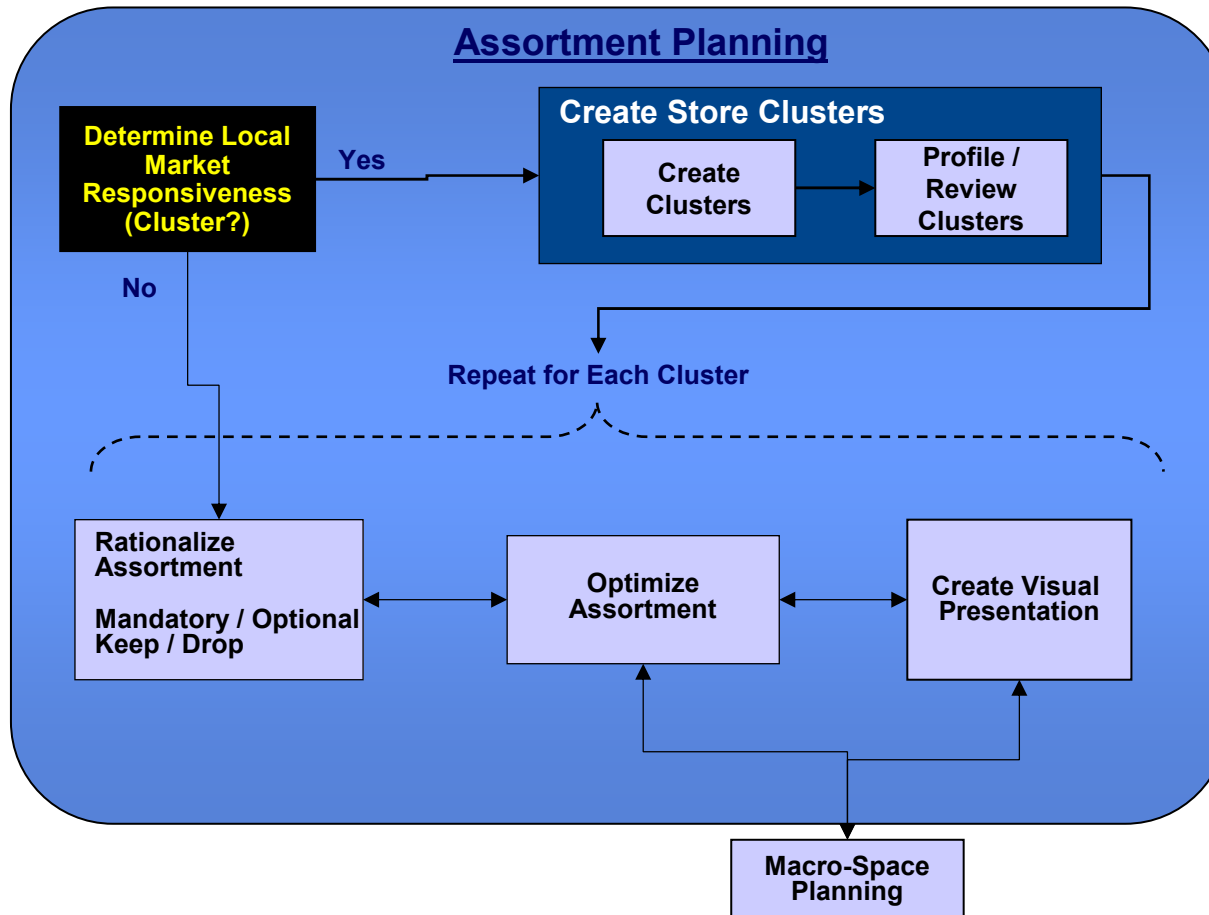
Product Availability is crucial for many retailers

Applications Accessed:

- ETL tool in Near Real time
- Teradata ADW
- On Shelf Availability Application
- BI Reporting Tool
- Existing Report Delivery System
- Stock Records POM System

- **Objective:**
  - > Improve on shelf availability of top 2000 skus in all stores

- **Analysis:**
  - > Identify top 2000 skus:
    - • Annual sales ranking (units, revenue, profit if available)
  - > Generate weekly rate of sale (units) matrix for each top sku for each store
    - • Use recent rolling 12 weeks history to generate ROS matrix
    - • Build in seasonality as DCM methodology
    - • Build in Promotion effects if available
    - • Use variable "time slice" (hour to multi-day) to reach minimum ROS
  - > Predict on shelf out of stocks before they occur within day for top 2000 goods

- **Actions:**
  - > Deliver reports/alerts to store personnel in "near real time"
  - > Check shelf / stockroom for stock
  - > Check on order / due date for next delivery if any / place order
  - > Check accuracy of stock records
  - > Rules based automatic ordering / replenishment direct to suppliers
  - > Improved "other income – rebates and allowances" from vendors based on in-stock position

- **Results:**
  - > >> 1% increase in sales
  - > Increased customer satisfaction
  - > Improved in-store procedures
  - > Improved stock accuracy
  - > Identification of Supply Chain Issues/Ownership

Teradata
a division of NCR

Developing a measure of on-shelf availability

- Objectives
  - > To develop a measure of product availability for a supermarket chain, that took account of low rates of sale

- Analysis
  - > Applied probability theory to model sales and identify periods where zero sales were most unlikely to have occurred by chance

- Action/Return
  - > Created a credible availability metric for use across all stores
 - > Reports were developed to track availability based on the probability model

Teradata
a division of NCR

Step 1: A probability model was fitted to the sales frequencies for fast and slow moving lines

Basket Frequencies - Fast Moving Line

Basket Frequencies - Slow Moving Line
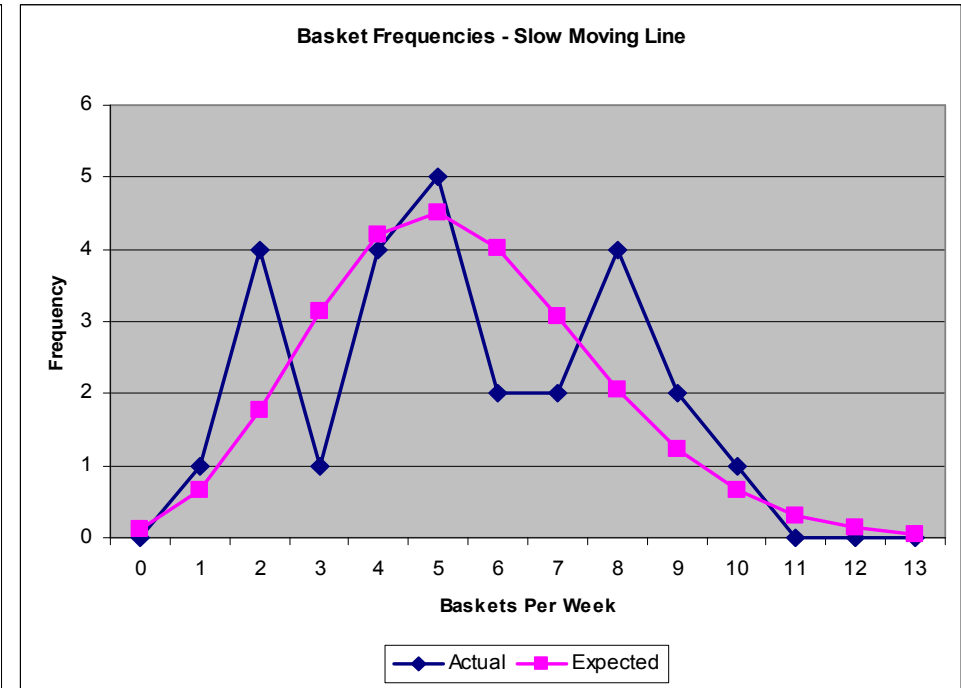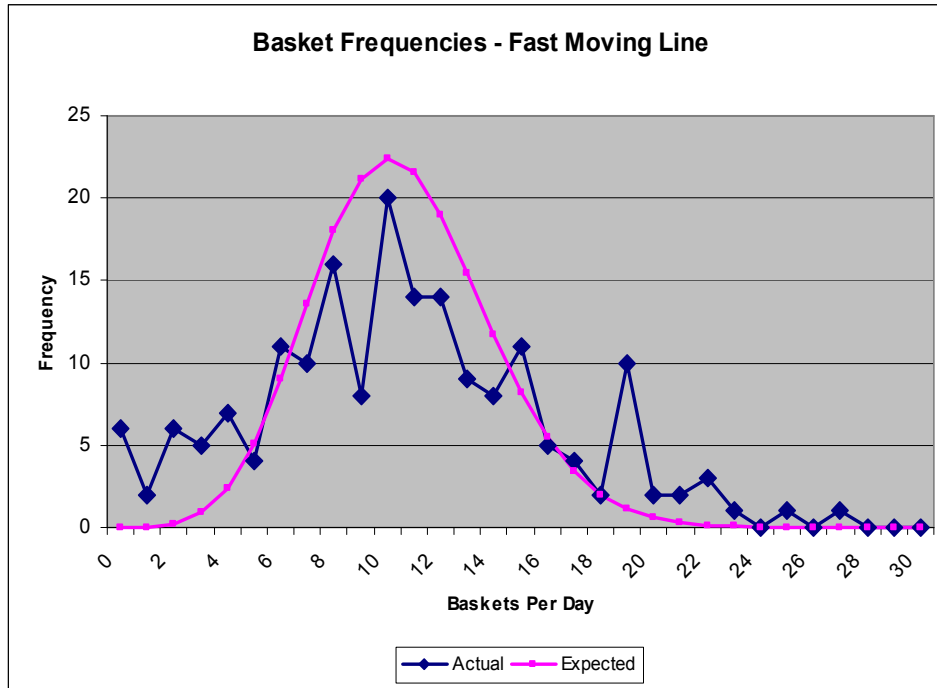
The model was sufficiently flexible to fit a wide range of products

# Step 2: Methods were developed for tracking fast and slow moving lines

Fast movers were tracked daily, lost sales were flagged and estimated for full and partial days

Daily Potential Lost Sales
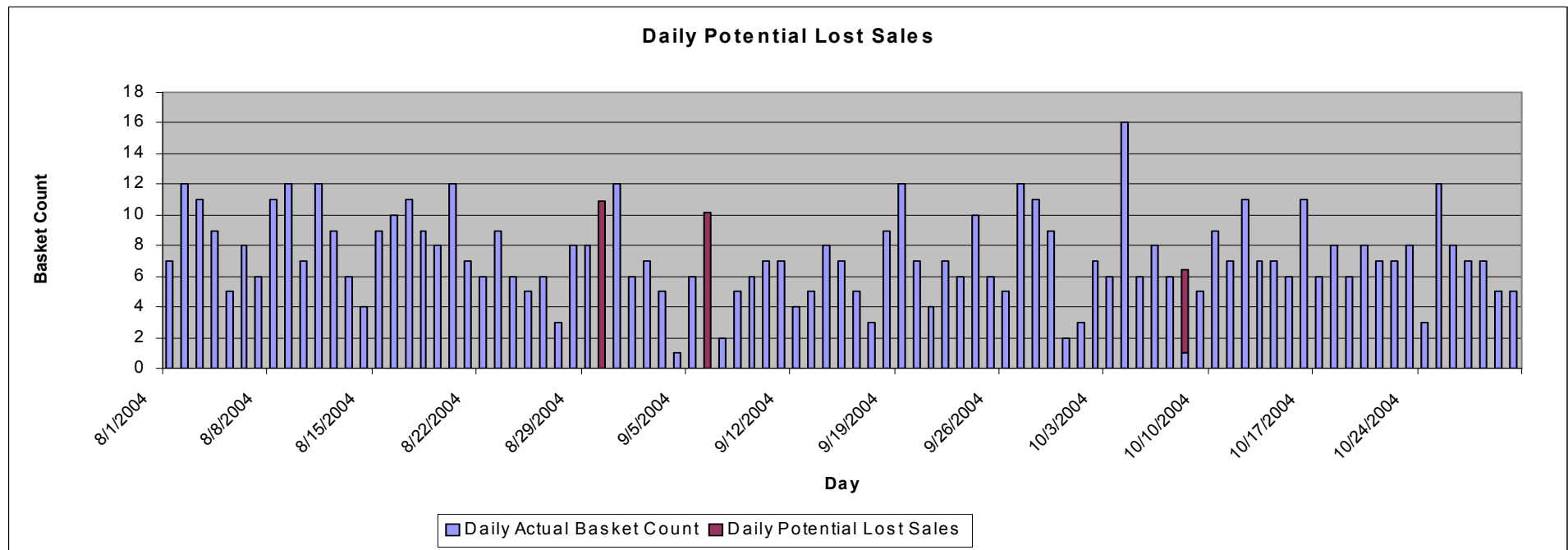
Example for a fast-moving line

Step 2: Methods were developed for tracking fast and slow moving lines

Slow movers were tracked on a rolling weekly basis, lost sales were flagged and estimated for full and partial weeks
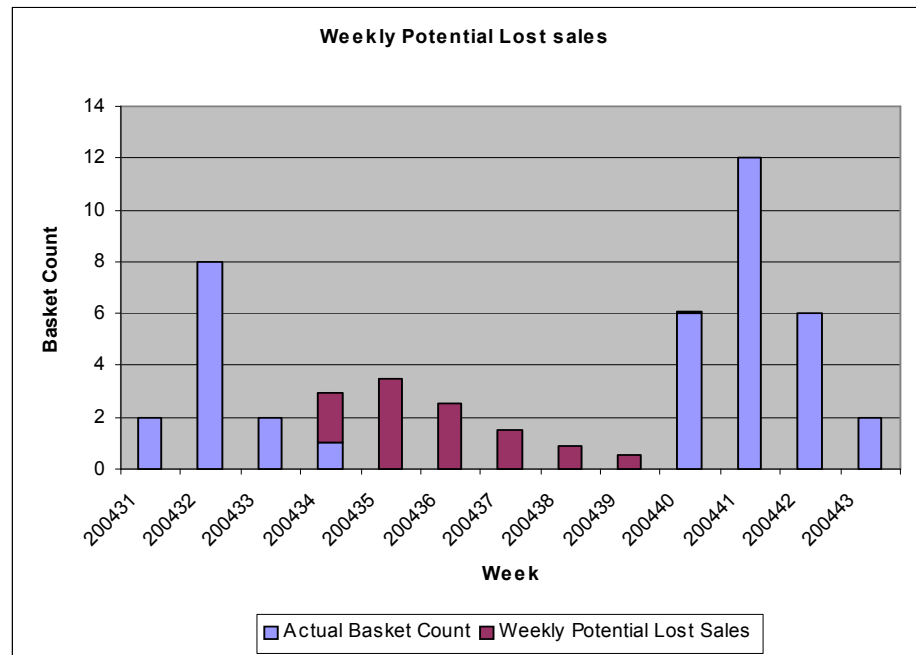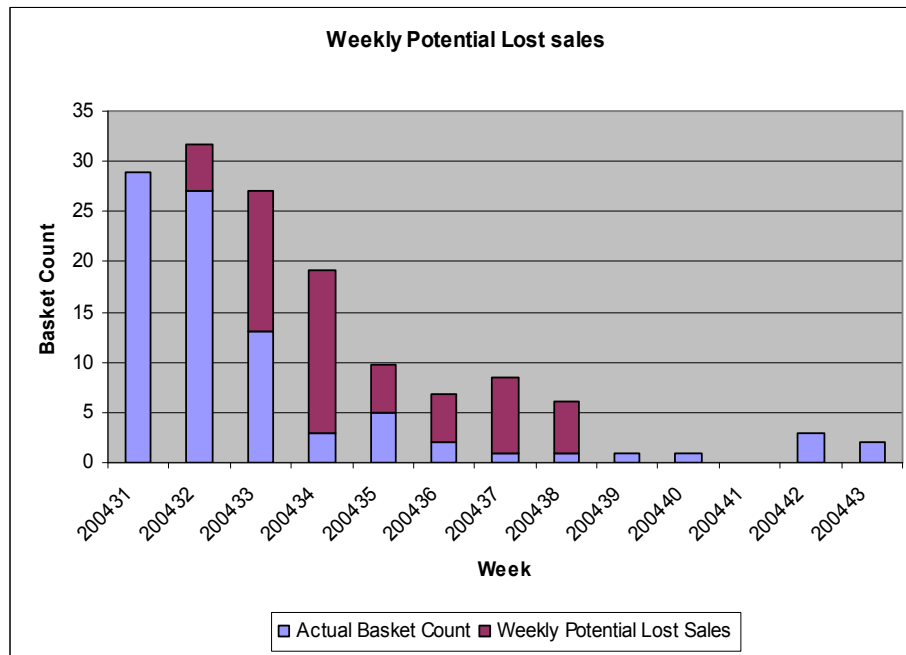
Examples for two slow-moving lines

Product Analytics - Discussion

- Have you applied these techniques?

- What are your views and comments?

Fraud Analytics - Overview

- Every industry suffers from fraud!
- Analytics have been very effective in fraud prevention and detection across all industries
  - > Underlying methods are largely the same
- For purposes of this session, we will categorize fraud detection analytics by level of sophistication and warehousing environment:

Sophistication

- **Basic – _anyone_ can implement them!**
- **Advanced Analytics**

Environment

- **Static – after the fact**
- **Active – catch them in the act!**

Teradata
a division of NCR

Methods Any User Can Apply (1)

- To detect activities <u>known</u> to match a previously **investigated suspicious pattern:**
  - \> Retail Industry Example:
    - Voided transactions around store closing
  - \> Telecommunications Industry Example:
    - Calls to numbers from previously-identified, fraudulent accounts
  - \> Financial Industry Example:
    - First-time credit-card holders making rapid/high volume purchases immediately upon receipt
  - \> Healthcare Industry Example:
    - Physicians providing services which are far outside their specialties
  - \> On-Line Auctions Example:
    - Shill bidding to artificially inflate the selling price

- These methods centre on building a routine to find matches with the pattern, and criteria to prioritize them for subsequent investigation
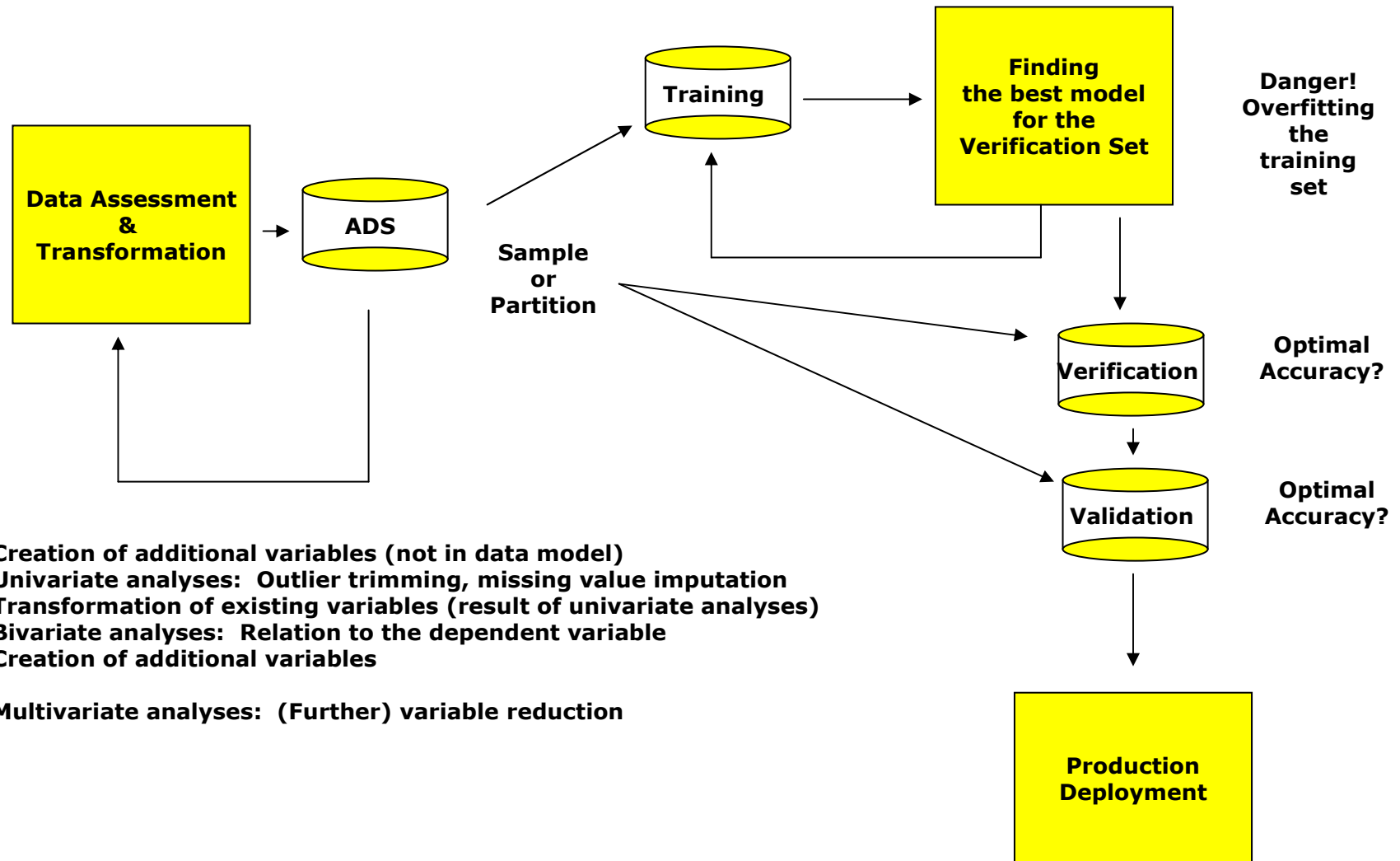
Methods Any User Can Apply (2)

- To detect patterns – not previously investigated - that <u>may</u> represent fraudulent activity

- These methods are based on the identification of abnormalities

- Such abnormalities exist in data as <u>*outliers*</u>, such as
  - > Sudden or dramatic changes in activity, relative to a population or industry profile
  - > Amounts whose size or frequency of occurrence is far from average

- These methods center on building a routine to find such outliers, and criteria to prioritize them for subsequent investigation

Advanced Fraud Detection Methods

- Sophistication stems from applying advanced analytic algorithms to fraud detection

- Algorithms may be "supervised", meaning known cases of fraud are provided to develop a predictive model
  - > Oversampling may be required to model relatively rare, fraudulent events
  - > Cost-benefit analysis should be applied to determine acceptable error types and rates associated with fraud detection models

- Or algorithms may be "unsupervised", meaning known cases are not provided
  - > As in the basic case, abnormalities are to be detected - but using advanced analytic algorithms such as clustering

- These methods centre on application of advanced analytics software, according to a data mining methodology e.g. CRISP-DM

Advanced Fraud Detection Methods

Data Assessment & Transformation

ADS

Sample or Partition

Training

Finding the best model for the Verification Set

Danger! Overfitting the training set

Verification

Optimal Accuracy?

Validation

Optimal Accuracy?

Creation of additional variables (not in data model)
Univariate analyses: Outlier trimming, missing value imputation
Transformation of existing variables (result of univariate analyses)
Bivariate analyses: Relation to the dependent variable
Creation of additional variables

Multivariate analyses: (Further) variable reduction

Production Deployment

Teradata
a division of NCR

A Propensity Model for On-Line Auction Fraud

BIDDER_ID	AUCTIONS PER WEEK	BIDS PER AUCTION	ALIASES USED	TIME BETWEEN BIDS	DIFFERENCE BETWEEN BIDS	AUCTIONS WON	ITEM TYPE INDICATOR	ITEM COST INDICATOR	DOMAIN TYPE INDICATOR	AVERAGE PRICE	FRAUDULENT
4497	18	3	7	45	1	13	3	1	470	2078	Yes
21054.72727	19	3	5	42	7	8	4	0	351	2150	No
9127.454545	20	4	3	20	4	8	8	0	378	767	No
20196.18182	23	2	7	24	4	11	7	1	267	507	Yes
14129.54545	19	2	5	35	5	9	5	0	239	1467	No
2402.818182	18	3	7	30	4	7	7	0	360	1670	Yes
17337.36364	17	2	5	22	4	6	7	0	275	3740	No
2413.909091	22	3	5	27	9	7	6	0	407	1666	Yes
10873.09091	31	3	4	34	8	6	17	0	586	1191	Yes
498.6363636	15	3	4	47	3	4	8	0	317	638	No

Can you spot the pattern that identifies certain on-line bidders as behaving fraudulently?

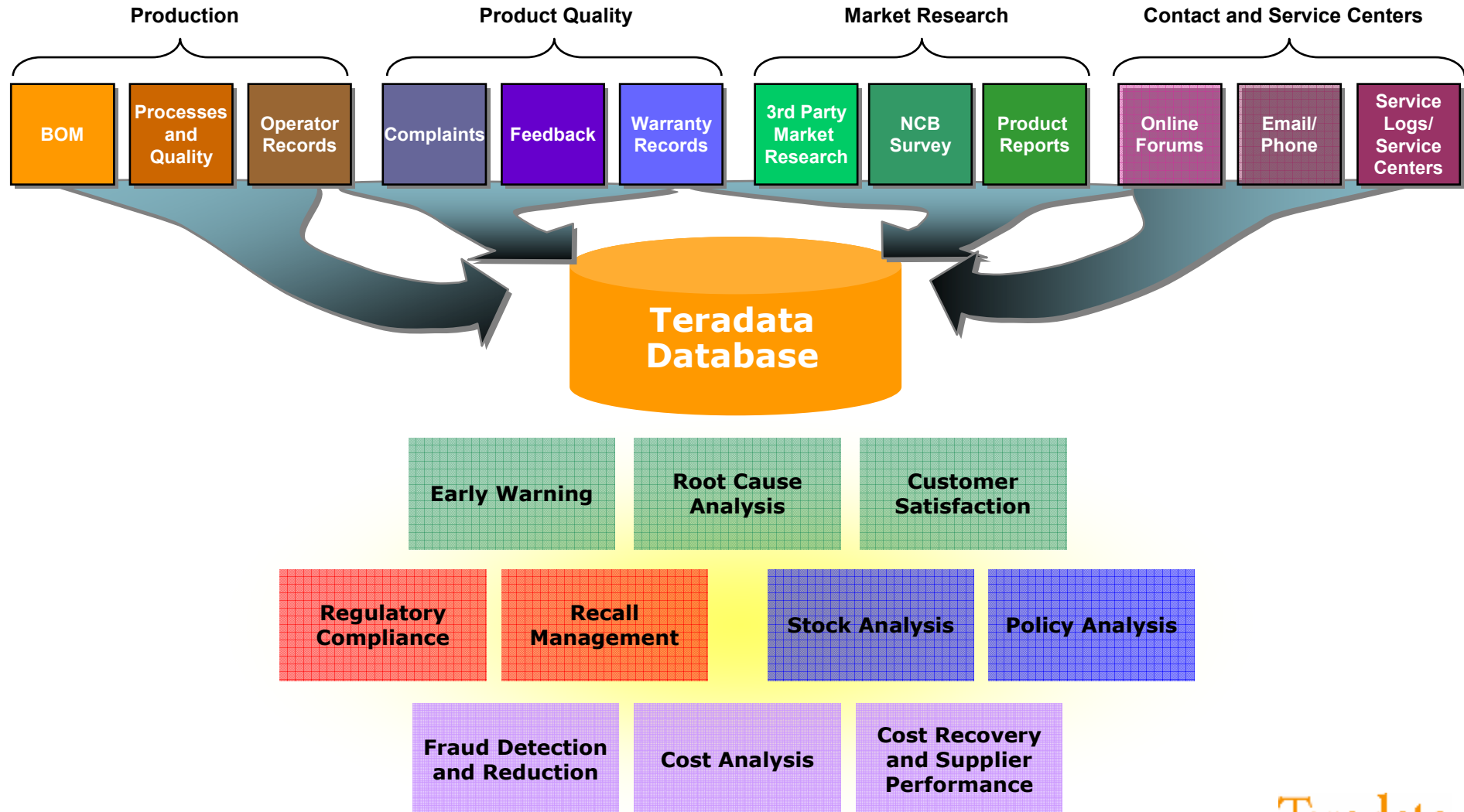A Decision Tree can quickly and accurately classify the members of this data set
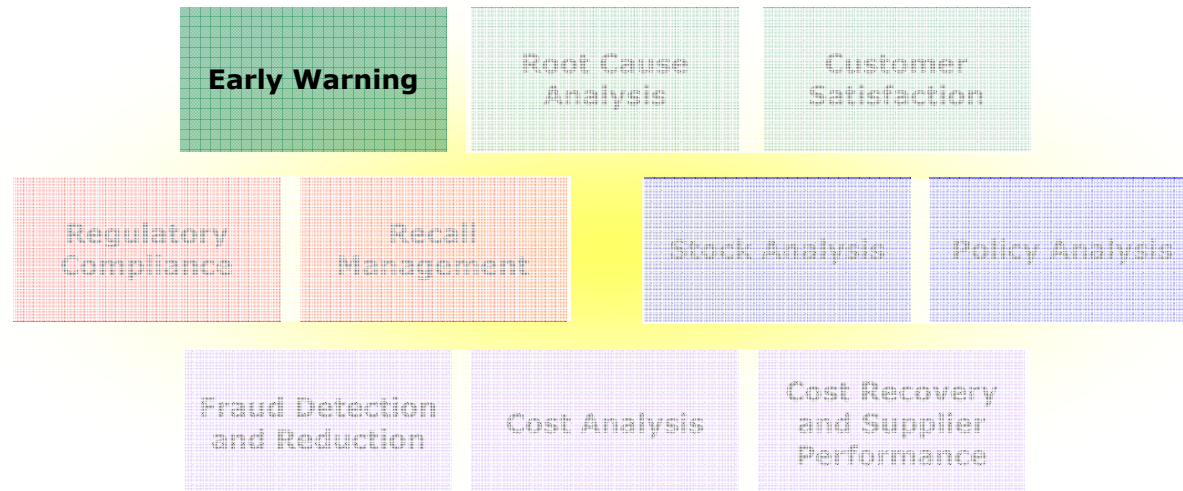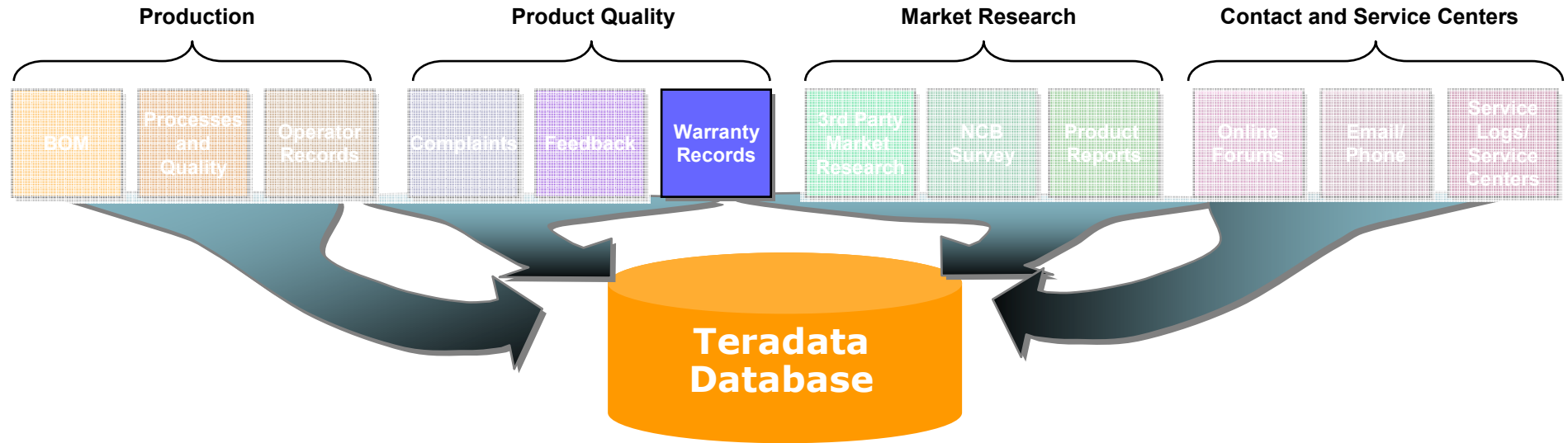
Decision Tree Results
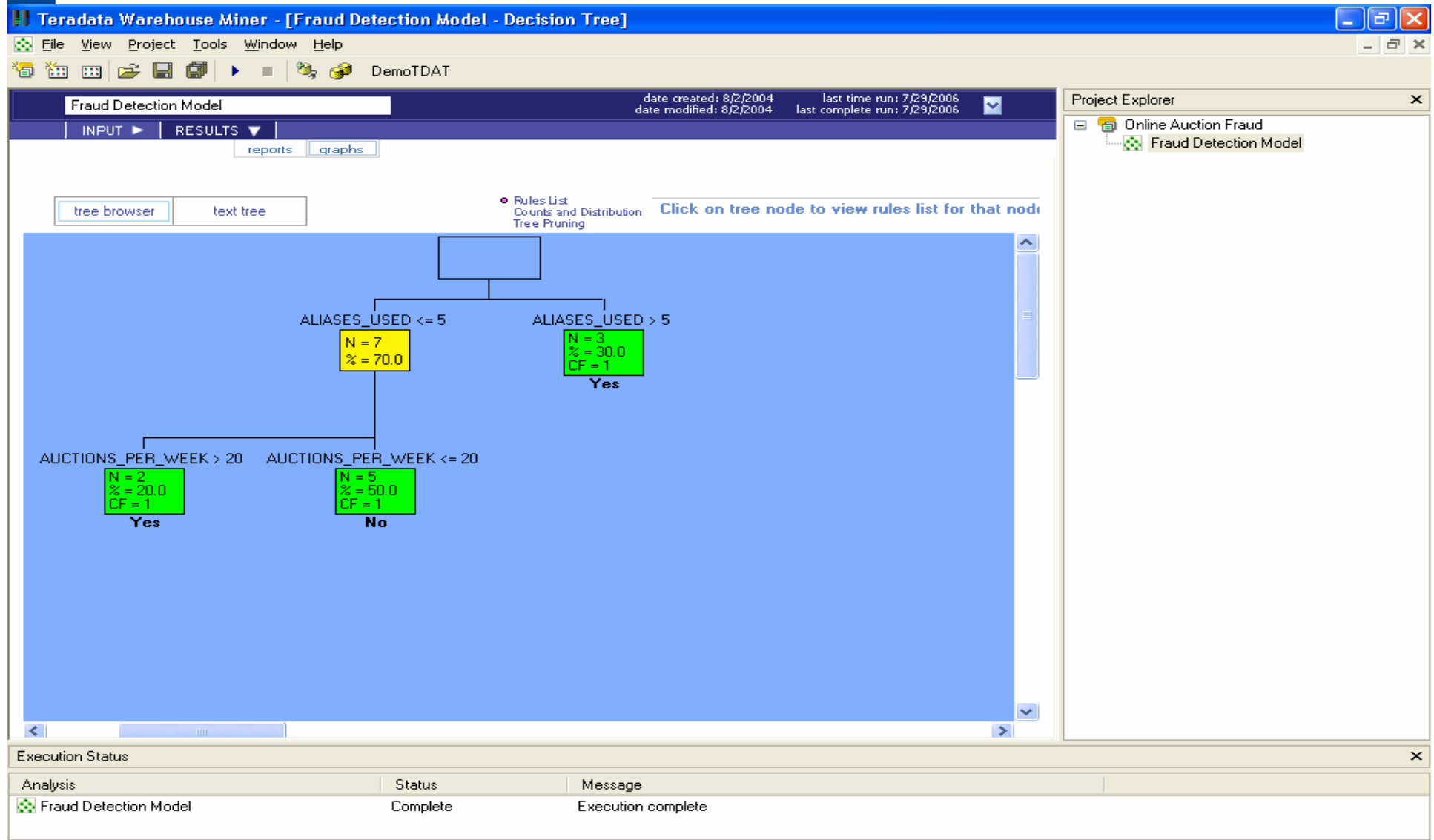
Fraud Analytics - Discussion

- Have you applied these techniques?

- What are your views and comments?

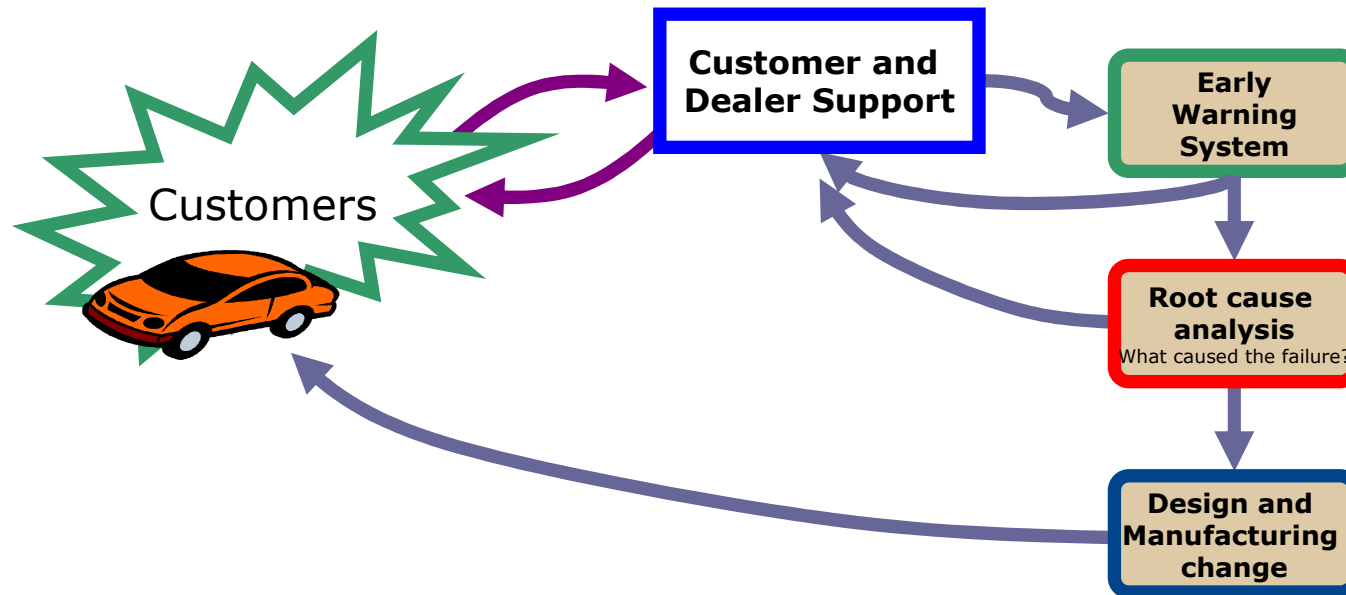Teradata
a division of NCR

Warranty Analytics:
Providing the framework for a enterprise approach
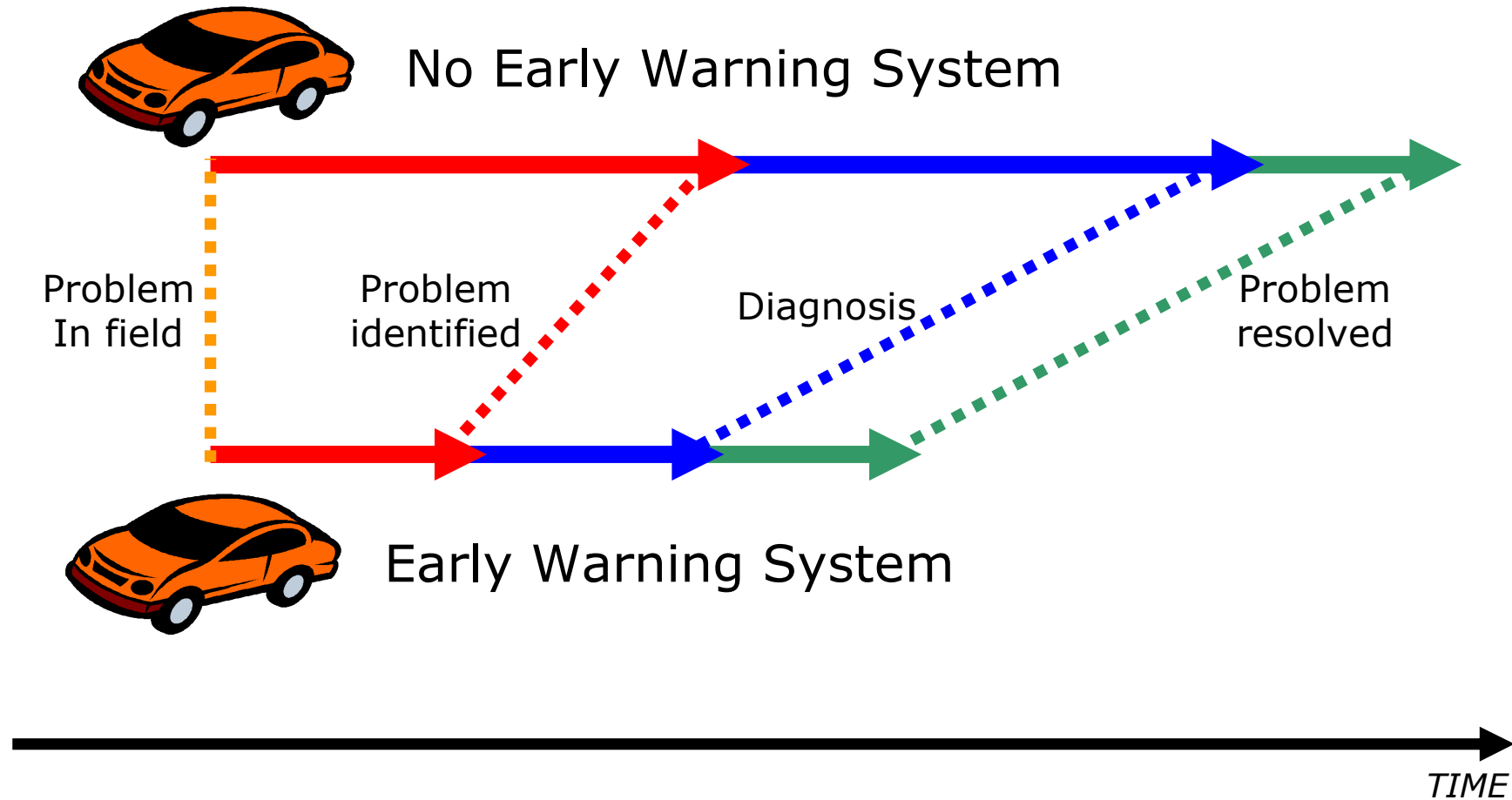
Step 1 - Early Warning Analysis

Quality and service for new vehicles: main objectives

- Top priorities for a manufacturer:
  - > improve the general satisfaction of new car buyers
  - > reduce the cash reserves for warranty costs

Reducing time to action

No Early Warning System

Early Warning System

Problem
In field

Problem
identified

Diagnosis
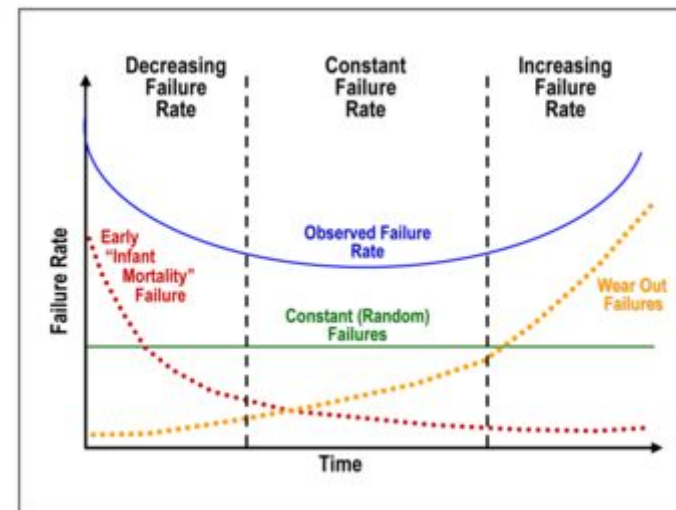
Problem
resolved

TIME

Teradata
a division of NCR

What do we mean by early warning system?

- ~~Text Analytics~~

- ~~Direct Read Out stuff~~

- ~~A bunch of reports~~

- A quantative approach to understanding how many failures there will be for a given component by a point in time
  - > Now we can understand which problems are large
  - > Root cause analysis can be prioritised
  - > Initially during launch period, but can be extended to whole lifecycle
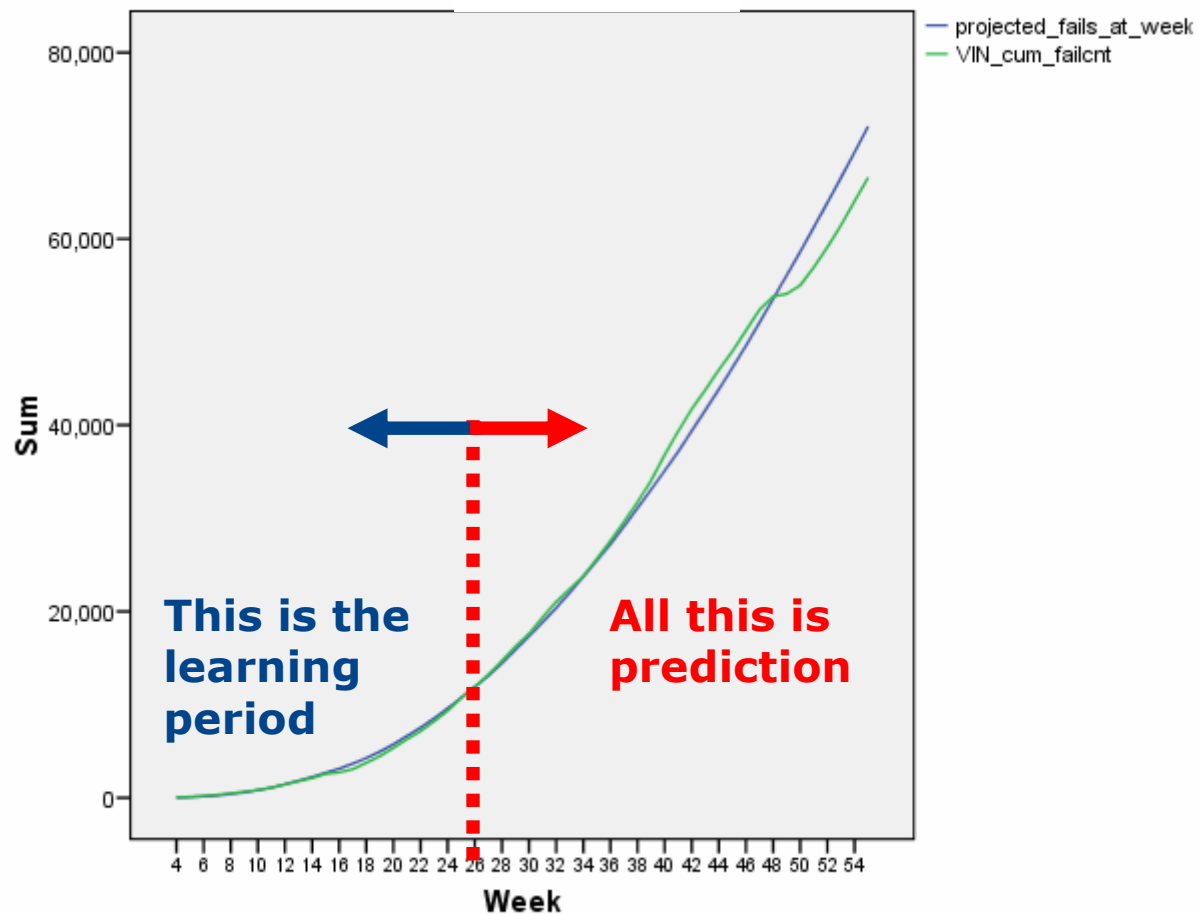
The theory

- Terminology:
  - > Hazard rate: the probability of a component failing in time t
  - > CDF: the probability of a component failing by time t

- We decided to use the Weibull function as an approximation of the CDF for all components

$$cdf = 1 - e^{-(t/\eta)^{\beta}}$$

- This approximates the CDF of a component following the bathtub curve

How do we evaluate the results?

- Example: 26 weeks of learning

The Early Warning System Solution

- The early warning system will enable the manufacturer to predict the size of problems early in the life of a model
  - > It can also be used to understand how problems develop over the life of the model

- It *learns* from existing failures, and *predicts* future failures
  - > Every item of the component hierarchy is analysed for every model
  - > Every week more data is available, and the system improves its learning

- Problems can be prioritised for action

Warranty Analytics - Discussion

- Have you applied these techniques?

- What are your views and comments?

Thank you for your participation!

Barry Leventhal
+44 7803 231870

You've never seen your business like this before.