Paper 180-28

# Rapid Analytic Application Deployment

David Press, Equitec, Cleveland, OH
Sam Iosevich, SAS, Chicago, IL
Larry Wasserman, SAS, Chicago, IL

## ABSTRACT

In the 21st Century, companies have to re-shape long-standing business practices and the rules of competition so as to address the rapidly changing marketplace.  They can no longer rely on tried-and-true methods or 'what has worked in the past' to maintain a competitive advantage in their marketplace.  These dynamics can be felt more no more directly than in the retail marketplace.

Retailers are now faced with the premise that their businesses are "over-saturated". The combination of a stagnant economy and too many undifferentiated entries into given retail markets have driven some retailers to the brink of extinction. Retailers around the globe share some common challenges.

- Over retailed market – saturation
- Lowered barrier to entry as a CPG competitor (offshore copycats )
- Retailers are protecting sales data (not sharing with suppliers)
- Decreasing customer loyalty / increasing expectations
- Desire to shift from mass advertising to targeted investments
- Diminishing returns from operations oriented initiatives
- Challenges with multi-channel customer focused initiatives
- Relatively few examples of integrated strategic customer knowledge
- Difficulty quantifying return on knowledge and technology investments

In the rapidly changing world of retail, businesses have to place bets using their capital to generate the largest, fastest, most predictable payback. To increase their level of confidence, they need the most accurate retail business intelligence to make decisions. Most of this intelligence should be derived around core markets, core customers, core merchandise, and their significant differentiators from national and local competition.

SAS/Equitec® have developed a framework to foster the integration of business intelligence initiatives across the enterprise.  The Rapid Analytic Application Deployment framework enables iterative adoption of various analytic techniques that deliver substantial business value.  The initial components of this framework deliver affordable and valuable information and provide the foundation for more advanced analytics that solve more complex business problems and support business strategy execution.

## INTRODUCTION

Businesses are looking to improve their ability to make informed decisions about where to invest to grow their business and where to cut cost without affecting their ability to deliver value to their valued customers. In addition, businesses are evermore focused on understanding their customers, their customers' needs, and how they are viewed within their own market. The thirst for information and quantified intelligence resultant from that information is changing the way strategic leaders are leveraging departments responsible for information technology management and statistical analysis. Some of the questions that retailers are pondering at the moment are:

- How do I derive value from my information?
- How should I position my company, brand and products in the market place?
- What do my customers look like and where are they located?
- How do my customers segment?
- What is my relative market potential?
- What is the competitive landscape in a given market?
- Why are my customers buying my product, or in a specific outlet?
- How do I optimize direct marketing activities?
- How can I eliminate customer confusion and best represent my brand positioning?

The need for in-depth intelligence and improved business agility is increasing the need for more instantaneous recall of information, immediate aggregation of information from different sources, and repeated analyses of business conditions and tactics. These dynamics, combined with less than satisfactory results from data warehousing initiatives, have opened the door for more innovative solutions in the areas of analytics and business intelligence.

The lifecycle of developing and executing analytic methodologies aligns with some of the pains associated with the development of technology solutions. Customer indecisiveness and misunderstanding creates project scope creep and inconsistent timelines. The complexity of data integration schemes to provide the optimum solution value proposition creates coordination problems. Solution developers share creative thoughts but not necessarily work products. Information technology development staffs have long grappled with these issues and have concluded that componentization and component re-use approaches provides a way to manage the unpredictability and "throw-away" idea generation.

## RAPID ANALYTIC APPLICATION FRAMEWORK

Consider a research/IT department that has been requested to provide a simple report summarizing a company's performance. "Simple" is a relative term.  Even if the report is summarizing the simplest metrics that are almost standard across all companies, the design and implementation of the report (e.g. defining metrics, required data, data accessibility, coding, data, report aesthetics, update cycle, job dependencies) can take months.

SAS/Equitec have developed a framework that is ready-to-use to eliminate the steps detailed above that requires slight data configuration while incorporating external data into the metrics. This rapid analytic application deployment framework can be Modified and improved upon to address different projects as they are developed.  Applying the framework enables the user to concentrate on content and metrics while providing a mechanism for re-use and a standard means for presentation of the results.

This framework is the first step in the process of delivering a knowledge-based solution.  Market basket analysis, customer profiles and scorecards are examples of analytic applications which will be further explored within this paper. As these applications become adopted by organizations as repeating
Fig. 1
practices, more sophisticated analytic applications can be built using the initial applications as base components. The adoption of these components will seed the desire for analytic solutions

| SKU | Vendor | SKU Description | Classification | SKU | Vendor | SKU Description | Classification | Correlation Level |
|---|---|---|---|---|---|---|---|---|
| 1111111 | Vendor A | Product Description | Product Type | 222222 | Vendor B | Product Description | Product Type | High |
| | | | | 333333 | Vendor B | Product Description | Product Type | High |
| | | | | 444444 | Vendor D | Product Description | Product Type | Medium |
| | | | | 555555 | Vendor E | Product Description | Product Type | Low |
| | | | | 666666 | Vendor F | Product Description | Product Type | Medium |

with increasing returns such as customer segmentation/valuation and advertising effectiveness. The re-use of these base components to assemble these higher-value applications will facilitate the emergence of analytic development standard practices and ultimately allow creative adaptations, which address the optimization of merchandising and inventory management.

### MARKET BASKET ANALYSIS

Market basket analysis provides suggestions for inventory optimization/merchandise mix which help retailers avoid nagging problems of overstock and clearance sales, while identifying any non-apparent cross-sell opportunities that might exist. This analysis is performed at one classification level above SKU and reported at a SKU level. A SKU level analysis is not usually feasible due to the granularity of the data and is usually not instructive because it makes a distinction at too fine a level. The analysis is designed to take seasonality and demography into account and is often an initial step toward a store assortment optimization initiative.

Using ODS within SAS the results of the market basket analysis are optimally displayed. (Fig. 1) The market basket analysis is automated using Proc Assoc within SAS/EM and is folded into macros administrated through SAS/WA.

### CUSTOMER PROFILES

Customer profiles are important for identifying populations to target with specific marketing efforts such as advertising and direct mail. The profile results might also suggest cross-sell opportunities if there are products that are readily identified with the targeted populations. For instance, if your customers tend to be higher-income, advertising efforts and store placement should include other high-end products.

Customer profiles are obtained using a combination of SAS analytics, client transaction data and Acxiom demographic data and data processes. The merging and cleansing of the data is performed via the AbiliTec™ process and made possible by assigning an AbiliTec link to each of the customers within the client transaction file. Rollups, purging and other data manipulation is performed very quickly through a process developed and managed by Acxiom. SAS analytics are then used to calculate univariate statistics as well as select the most significant variables for profiling. Variables are chosen on their ability to distinguish the top 25% of the customers from the bottom 75% (Fig. 2). This process is again automated via SAS macros, ODS and Warehouse Administrator. Once the client's data is mapped into the rapid analytics application deployment framework warehouse, the profiling process is instantaneous and can be deployed anywhere via SAS/EIS.
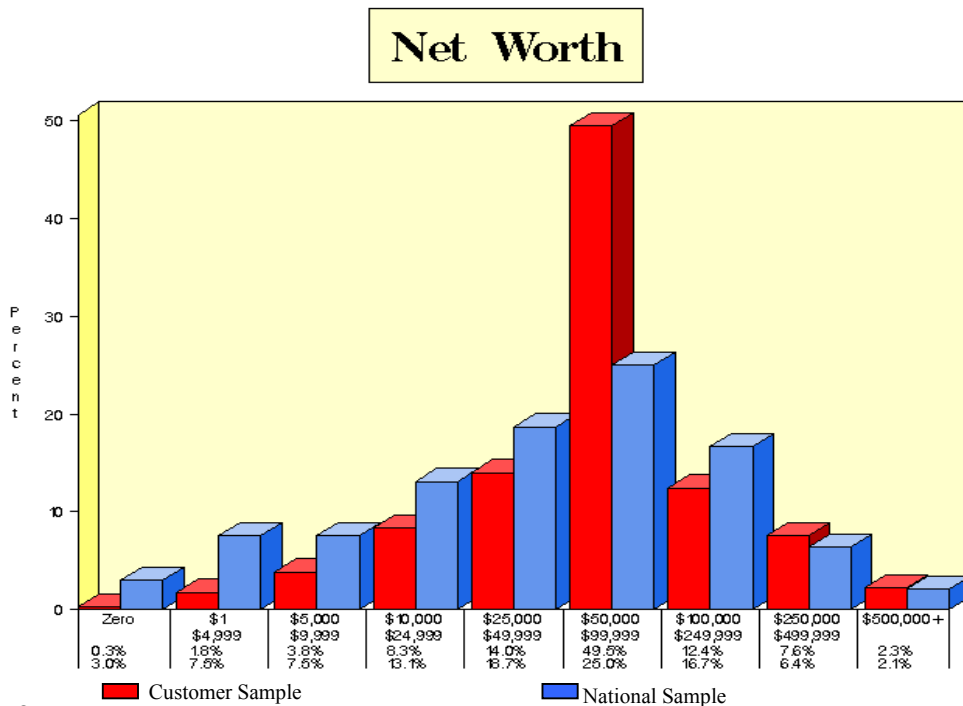


Fig. 2

**EQUITEC®**

**\<Company\> Scorecard (7/1/01 - 6/30/02)**

| Retailer Name | Store Number | Store Location | Current Performance Index | Potential Index | Consumer Index | Professional Index | Competitive Index |
|---|---|---|---|---|---|---|---|
| Retailer A | 101 | QUINCY, MA | 202 | 504 | 358 | 727 | 448 |
| Retailer A | 102 | BROOKLYN, NY | 202 | 362 | 360 | 280 | 240 |
| Retailer A | 103 | NORTH MIAMI BEACH, FL | 201 | 0 | 22 | 27 | 44 |
| Retailer A | 104 | MANCHESTER, NH | 200 | 68 | 100 | 98 | 109 |
| Retailer A | 105 | SILVER SPRING, MD | 200 | 346 | 279 | 456 | 306 |
| Retailer A | 106 | PHILADELPHIA, PA | 200 | 270 | 283 | 207 | 186 |
| Retailer A | 107 | LONG ISLAND CITY, NY | 200 | 0 | 35 | 36 | 66 |
| Retailer A | 108 | STATEN ISLAND, NY | 199 | 0 | 339 | 227 | 164 |
| Retailer A | 109 | GAITHERSBURG, MD | 197 | 68 | 89 | 115 | 120 |

Fig. 3

**SCORECARDS**

Scorecards allow a quick and simple way to compile this information into an easy-to-reference "dashboard" reporting tool, which provides fundamental information organized by distribution channels and outlets. The metrics themselves are useful in that they show where the customers and competitors are, but it's the combination of metrics that provide the most insight because they get at potential: they help identify those stores that are doing better than expected and those that need something to push them up to their potential (consumers and professionals are there, but the sales aren't – why not?). This information is made more valuable by breaking it down by geographic trade area, which makes the information more actionable, and is at a level at that previously has not been readily available.

Using SAS analytics, it is determined which stores have the greatest potential for improvement based on the demographics, competitive landscape and sales performance of an area. This is derived using the customer profiles generated. The demographic drivers are found using SAS/EM decision trees. The optimal splits within the variables are also obtained. The Key Performance Indicators (KPIs) are compiled and represented via SAS (fig. 3). The process is once again completely automated up to the client data.

The highly automated characteristic of the analysis allows Equitec and SAS to deliver a quick and inexpensive solution, which can then justify a more in depth treatment of the client's business challenges. The data compilation and analysis is performed in 3-4 weeks. Once the cause of the business challenges has been determined through the rapid analytic application deployment framework, other solutions such as store assortment optimization may be performed. These solutions are easily financed with the ROI obtained from execution of business intelligence initiatives this framework facilitates. The results from this framework are updated on a quarterly basis to facilitate an early detection system for the client's business.

**THE FOUNDATION:  UNDERLYING DATA STRUCTURE**

The rapid analytic application deployment framework relies on an underlying data structure.  The relevant data is mapped into this structure. The data map and the underlying data models were envisaged to encourage and permit a standard and flexible data processing structure for the required analytic output.  The basic data map as explained in the diagram in figure 4 was composed of two major sections.  The basic ETL process takes the customer "input" files and loads them into the data warehouse. These warehouse tables are subsequently used as input to the data mart process used to create the required analytic data marts and associated reports.

The actual analytic data model portion for producing various reports focused on a set of data marts for each of the required analytic components.   This resulting data process assumes that the source tables have a consistent naming structure so that one data process can be used for all customer submissions.  One such data mart process referred to as the "customer profile" is shown in figure 5.
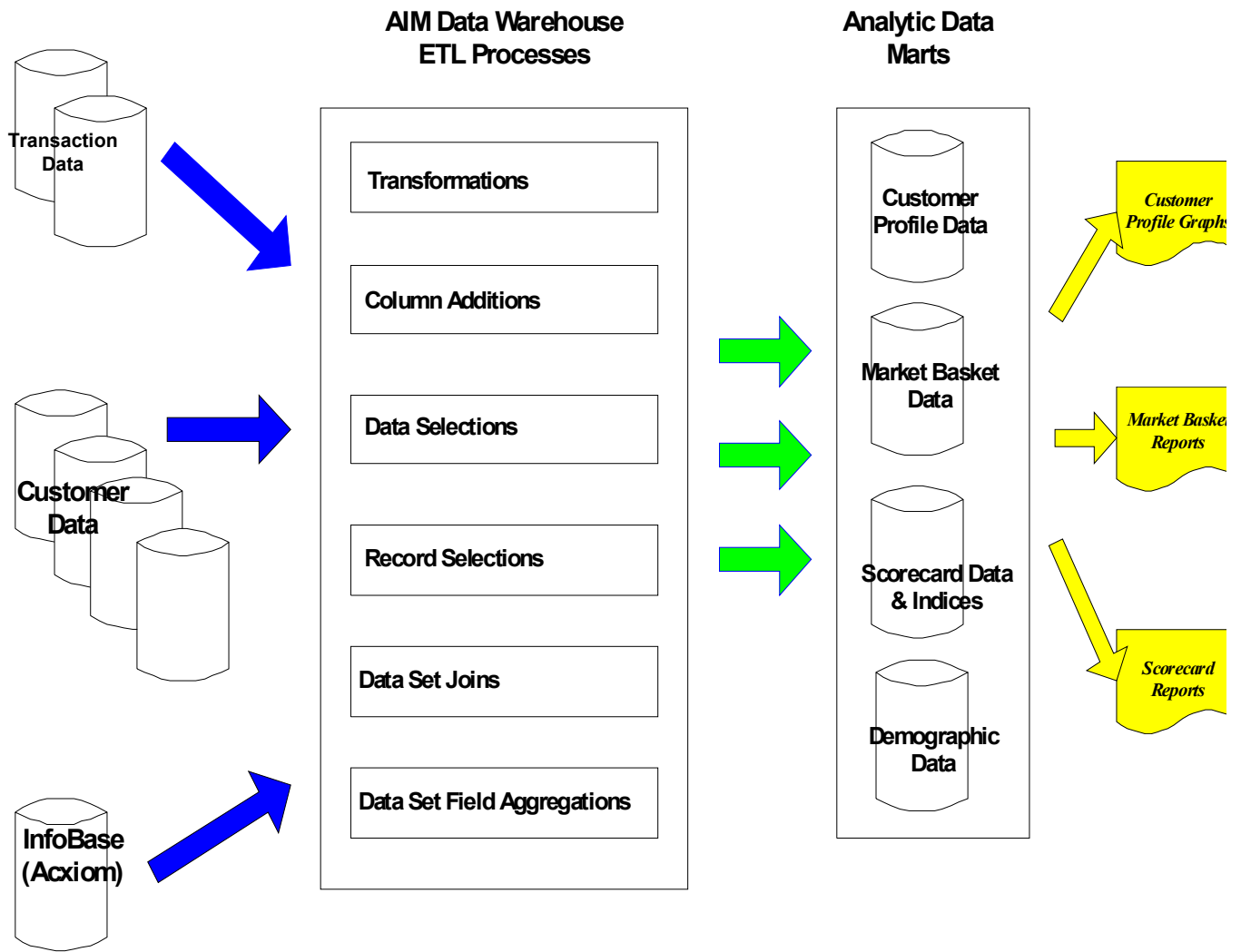
**AIM Data Warehouse
ETL Processes**

**Analytic Data
Marts**

Transaction
Data

Transformations

Column Additions

Data Selections

Record Selections

Data Set Joins

Data Set Field Aggregations

Customer
Data

InfoBase
(Acxiom)

Customer
Profile Data

Market Basket
Data

Scorecard Data
& Indices

Demographic
Data

*Customer
Profile Graphs*

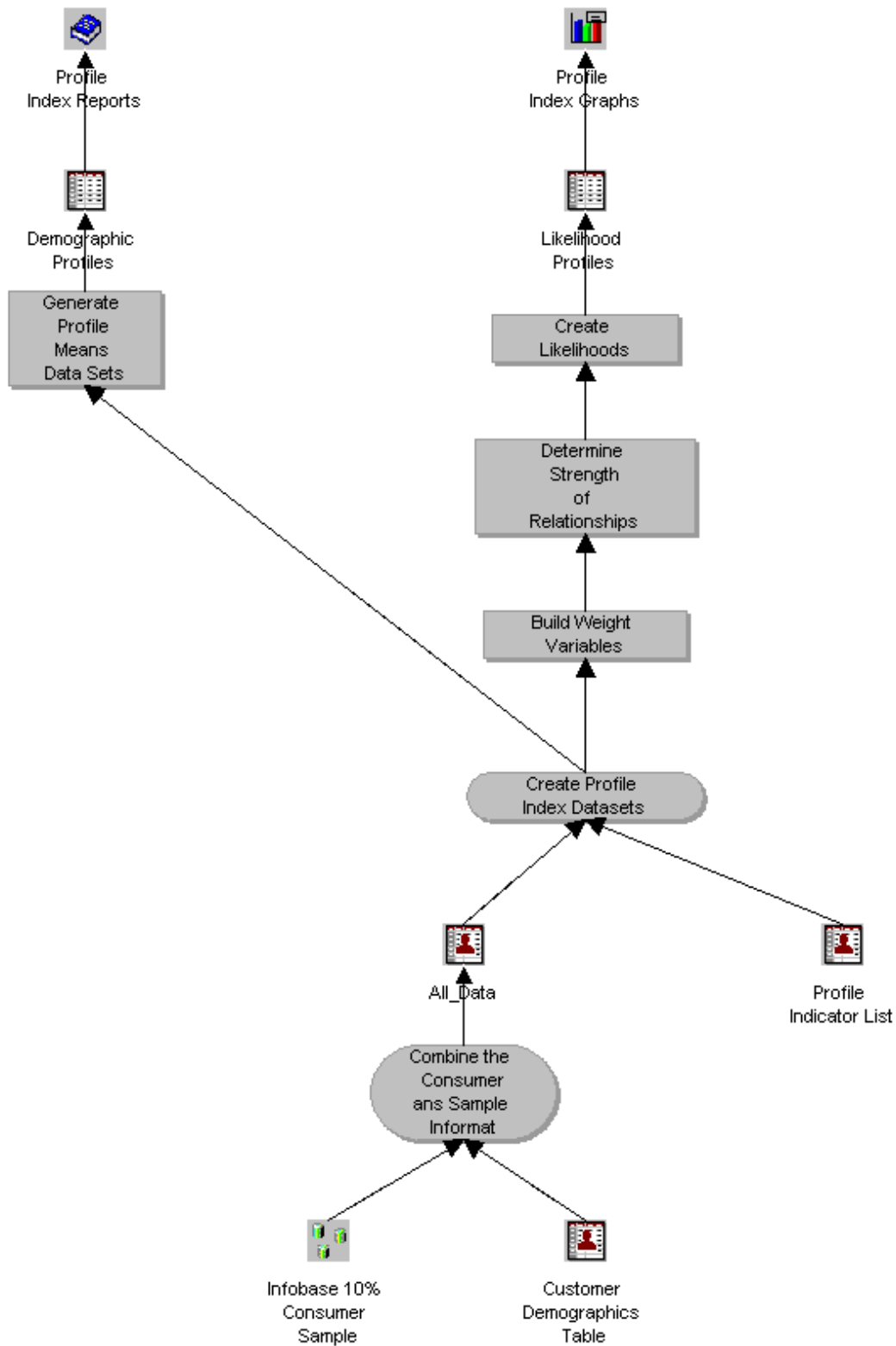*Market Basket
Reports*

*Scorecard
Reports*

Fig. 4

Fig. 5

## CONCLUSION

In conclusion, there is a compelling argument to bring rapid application development principals to the analytic development world.  SAS and SAS software solutions are key components to this effort, and SAS customers will benefit from frameworks such as those described within.

## CONTACT INFORMATION

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

> David Press
> Equitec
> 905 Corporate Way
> Westlake, OH  44145
> Tel: 216.881.2594
> Fax: 216.881.2589
> Email: dpress@equitec.com
> Web: www.equitec.com

> Sam Iosevich, Larry Wasserman
> SAS Institute Inc.
> 2 Prudential Plaza
> 180 North Stetson St., Suite 5200
> Chicago, IL  60601
> Tel: 312.819.6800
> Fax: 312.240.0342
> Email: sam.iosevich@sas.com,
> larry.wasserman@sas.com
> Web: www.sas.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Equitec is a registered trademark of The Henry Group, LTD. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.